Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 14, No. 3, 2026, 275-289.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

Predicting Hydrogen Combustion Characteristics Under Nitrogen Dilution: A Comparative Evaluation of GPR, MLP, and DNN Models

Ali Asadi[1] and Morteza Noferesti[2*]

1. Department of Mechanical Engineering, Bozorgmehr University of Qaenat, Qaen, Iran.
2. Department of Computer Engineering, Bozorgmehr University of Qaenat, Qaen, Iran.

Article Info

Article History:
Received 13 June 2025
Revised 26 February 2026
Accepted 05 July 2026

DOI:10.22044/jadm.2026.16390.2761

Keywords:
Deep Learning, Fuel Dilution, Hydrogen Combustion, Machine Learning, NOx Emissions, Perfectly Stirred Reactor.

**Corresponding author: mnoferesti@buqaen.ac.ir (M. Noferesti).*

Abstract

Hydrogen combustion has emerged as a pivotal technology for decarbonizing the energy sector, offering a clean and sustainable alternative to fossil fuels. This study investigates hydrogen combustion dynamics in a perfectly stirred reactor (PSR) under steady-state, non-premixed conditions. It employs CHEMKIN-based simulations to analyze the effects of nitrogen dilution, operating pressure, and equivalence ratio on flame temperature and NOx emissions. The parametric results reveal that nitrogen dilution reduces flame temperature by up to 28% and suppresses NOx emissions by 15–40%, while elevated pressure promotes higher flame temperatures and increased NOx formation. Peak temperature and NOx concentrations are observed under stoichiometric conditions ($\varphi = 1.0$), with both quantities decreasing under lean and rich mixture conditions. To enable rapid and accurate prediction of these combustion characteristics, three machine learning models were developed and benchmarked on the CHEMKIN-generated dataset: Gaussian Process Regression, Multilayer Perceptron, and Deep Neural Network. GPR demonstrated the best overall predictive performance, achieving the lowest MAE for temperature prediction (MAE = 1.77) and major species concentrations. Although the DNN produced competitive results (MAE = 58.28), it demanded approximately five times more computational resources than GPR, without a proportional gain in predictive accuracy. All three models maintained prediction errors below 20% across the investigated parameter space, confirming their viability as efficient tools for hydrogen combustion optimization. These findings demonstrate that physics-informed machine learning models, when combined with high-fidelity combustion simulations, offer a powerful and computationally efficient pathway toward accelerating the design and optimization of clean hydrogen energy systems.

1. Introduction

Hydrogen is increasingly recognized as a key solution for reducing global greenhouse gas emissions, with demand for hydrogen and hydrogen-based fuels expected to rise significantly in the coming decades [1]. This shift is critical, as fossil fuel-based energy currently accounts for approximately 60% of global greenhouse gas emissions [1]. Transitioning to low- or zero-carbon energy sources, such as hydrogen, can mitigate the release of harmful pollutants, including carbon dioxide, carbon monoxide, and unburned hydrocarbons. However, hydrogen combustion produces nitrogen oxides (NOx) as a byproduct, necessitating strategies to minimize their formation.

Diluted combustion, achieved by introducing inert gases like nitrogen, is one such approach that reduces NOx emissions by lowering flame temperatures and altering chemical pathways [2–

7]. The dilution of hydrogen flames thickens the flame structure and reduces NOx emissions due to lower local adiabatic temperatures and chemical effects [2]. The equivalence ratio further influences thermal and chemical effects, as demonstrated in studies of one-dimensional free flames [5].

Numerical simulations of non-premixed hydrogen combustion using Large Eddy Simulation have revealed that the equivalence ratio in the flame zone is often leaner than the designed value, with NOx formation mechanisms varying across flame, post-flame, and recirculation zones [8]. A study on laminar hydrogen flames in a counterflow model revealed that thermal NOx is significant at high temperatures, while NO formation through N2O and NNH pathways becomes notable at lower temperatures [9]. The dominance of thermal NOx and the importance of N2O and NNH pathways have also been reported in other studies on hydrogen combustion. Among the NNH-involved reactions investigated by [10], it was found that NO is highly sensitive to the reaction of NNH with the O radical, NNH+O NH+NO. The significance of this reaction in NO production has been clearly understood since 1995 [11]. Recently, Brier et al. [12] also observed that NO formation through the NNH pathway is one of the key mechanisms in the flame zone for pure hydrogen combustion, while thermal, extended Zeldovich, and N2O pathways are more prominent in the post-flame region.

Hydrogen combustion is characterized by unique challenges, including high flame speeds, elevated adiabatic flame temperatures, and the emergence of thermoacoustic instabilities [13,14]. These complexities demand advanced modeling approaches, particularly for multi-scale simulations that account for both chemical reactions and acoustic interactions. AI-based techniques, such as multilayer perceptrons (MLPs), long short-term memory (LSTM) networks, and Gaussian Process Regression (GPR), have proven effective in predicting nonlinear flame dynamics and refining flame descriptor functions for hydrogen-enriched flames [15-17]. By enhancing predictive accuracy while reducing computational costs and time, these methods have become essential tools for advancing research and practical applications in hydrogen combustion systems.

In this paper, we propose a hybrid computational approach that integrates high-level combustion characteristics with advanced machine learning techniques to enable rapid and accurate prediction of hydrogen combustion dynamics. The proposed approach bridges the gap between detailed chemical kinetics and data-driven modeling by evaluating three machine learning models Gaussian Process Regression (GPR), Multilayer Perceptron (MLP), and Deep Neural Network (DNN) for predicting key combustion characteristics under varying dilution levels and pressure conditions. Our methodology enhances the ability of machine learning models with CHEMKIN-based numerical simulations to: (1) predict the effects of nitrogen dilution and equivalence ratio on flame temperature and NOx formation, (2) establish performance benchmarks for GPR, MLP, and DNN models in reproducing complex combustion phenomena, and (3) develop predictive tools for industrial hydrogen system design.

The main contributions of the paper are as follows:

1. A hybrid approach integrating CHEMKIN-based numerical simulations with three machine learning models (GPR, MLP, and DNN) to enable rapid and accurate prediction of hydrogen combustion dynamics under varying dilution levels and pressure conditions.
2. Systematic evaluation and performance benchmarking of GPR, MLP, and DNN models for predicting flame temperature and NOx formation as a function of nitrogen dilution and equivalence ratio.
3. Development of predictive tools that advance fundamental understanding of hydrogen combustion and accelerate the design of industrial hydrogen energy systems.

The rest of the paper is organized as follows: Section 2 reviews related work on hydrogen combustion modeling and machine learning applications in energy systems. Section 3 details the theoretical foundations and numerical modeling approach, including reactor configuration and chemical kinetics. Section 4 introduces the machine learning methodologies and their implementation. Section 5 presents a comprehensive evaluation of model performance against simulation data, analyzing accuracy, computational efficiency, and robustness. Finally, Section 6 concludes with key findings and outlines future research directions for AI-augmented combustion design.

2. Related Works

Recent advances in artificial intelligence and machine learning have influenced combustion research by providing computationally efficient alternatives to conventional numerical simulations and experimental investigations. One research direction has focused on predicting flame dynamics and thermoacoustic behavior. In this context, [18] investigated the capability of machine learning

algorithms to predict the nonlinear response of premixed laminar flames subjected to acoustic excitation. Their results showed that multilayer perceptron networks outperformed Gaussian Process Regression and Random Forest models, particularly when combined with physics-based preprocessing techniques. Although the incorporation of flame response time improved extrapolation performance, the prediction accuracy at higher hydrogen enrichment levels indicated limited robustness when operating conditions deviate significantly from the training data. This limitation highlights the need for models that maintain predictive accuracy across wider parameter ranges.

A second group of studies has concentrated on predicting combustion emissions and engine performance. Mao et al. [19] developed artificial neural network models for estimating NOx emissions in ammonia–hydrogen combustion systems. Their work demonstrated that ANN-based approaches can effectively replace computationally expensive chemical reaction network simulations. Ramalingam et al. [20] employed ANNs to predict the performance of diesel engines operating with ammonia-, hydrogen-, and biodiesel-based fuel blends, achieving high prediction accuracy. Studies by [21] and [22] further utilized ANN models to optimize hydrogen–diesel and biodiesel-fueled engine configurations, respectively, showing substantial improvements in performance and emission characteristics. While these investigations confirm the effectiveness of data-driven methods for regression and optimization tasks, they primarily rely on black-box learning frameworks that provide limited physical interpretability and often require large quantities of high-quality training data.

Another active area of research involves the direct prediction of combustion-related parameters using machine learning. For example, [23] proposed several ML models for estimating turbulent flame speed in spark-ignition engines fueled by hydrogen–natural gas mixtures. Their results demonstrated substantial computational savings compared with conventional approaches, with Random Forest models yielding the best overall performance. Nevertheless, the study focused on a specific combustion characteristic and did not explicitly incorporate governing physical principles into the learning process, which may restrict model reliability under unseen operating conditions.

Physics-Informed Neural Networks (PINNs) have emerged as a promising paradigm for solving partial differential equations and modeling complex physical systems by embedding conservation laws and governing equations directly into the training process. In combustion applications, PINNs have been applied to flame dynamics, reactive flow modeling, and combustion system identification, offering improved physical consistency and enhanced generalization compared with conventional neural networks. Similarly, reduced-order combustion modeling techniques have gained considerable attention as efficient surrogates for high-fidelity computational fluid dynamics simulations, enabling accurate representation of combustion behavior with significantly reduced computational requirements. These approaches address one of the principal shortcomings of purely data-driven models: their limited ability to extrapolate beyond the range of available training data while preserving physical realism.

Despite growing interest in AI-based combustion modeling, significant gaps remain in the literature. Studies on premixed or turbulent flame dynamics [15–18] rely on simplified flow assumptions that overlook the kinetic sensitivities of hydrogen flames, while engine-focused studies [19–23] prioritize macro-level performance metrics over detailed thermochemical interactions governing NOx formation and flame temperature. Furthermore, prior machine learning approaches have employed single-model frameworks without comparison and lack physics-informed feature engineering tailored to hydrogen-specific kinetics. In contrast, the present work proposes a hybrid physics-informed machine learning framework targeting PSR-based non-premixed hydrogen combustion, integrating CHEMKIN-generated simulation data with engineered input features including log-pressure scaling, N_2/O_2 ratio, and kinetic interaction terms, to represent the underlying combustion physics. Unlike previous studies, GPR, MLP, and DNN models are conducted on a unified dataset with identical preprocessing and evaluation metrics, providing a comparative assessment of hydrogen flame temperature and NOx emissions under varying nitrogen dilution levels and equivalence ratios.

3. Theory and Modeling

The perfectly stirred reactor, also known as the ideal well-stirred reactor, is an idealized reactor model where complete mixing is achieved within the control volume. A simplified schematic of a perfectly stirred reactor is illustrated in Figure 1. Experimental reactors that utilize high-speed inlet jets closely approximate this ideal condition and have been widely used to study various aspects of

combustion, such as flame stability and NOx formation [24].

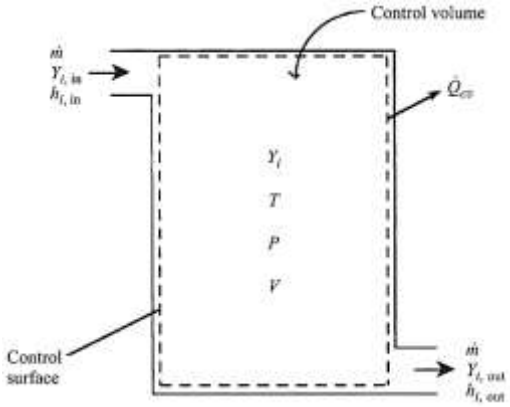

Figure 1. Simplified Schematic of a Perfectly Stirred Reactor (Turns, 2000).

The conservation of mass for an arbitrary species *i* within the control volume of the perfectly stirred reactor (Figure 1) can be expressed as follows:

$$\frac{dm_{i,cv}}{dt} = \dot{m}_i''' \cdot V + \dot{m}_{i,in} - \dot{m}_{i,out} \quad (1)$$

In this equation, m, $\dot{m}$, t, and V represent the mass of the control volume, mass flow rate, time, and volume, respectively. The term $\dot{m}'''$ corresponds to the production or consumption of species *i*. When the appropriate form of Equation (1) is written for each species in the reactor ($i = 1, 2, \ldots, N$), the sum of these equations yields the familiar form of the continuity equation (Equation 2).

$$\frac{dm_{cv}}{dt} = \dot{m}_{in} - \dot{m}_{out} \quad (2)$$

The mass production rate of a species is related to the net production rate $\dot{\omega}_i$ through the following relationship:

$$\dot{m}_i''' = \dot{\omega}_i MW_i \quad (3)$$

where MW_i represents the molecular weight of species *i*. By neglecting any diffusive flow, the mass flow rate of each species can be simply obtained by multiplying the total mass flow rate by the mass fraction of that species, Y_i :

$$\dot{m}_i = \dot{m} \cdot Y_i \quad (4)$$

If we apply Equation (1) to a perfectly stirred reactor and assume steady-state conditions, the time derivative on the left-hand side vanishes. With this assumption and by substituting Equations (3) and (4), Equation (1) simplifies to the following form for $i = 1, 2, \ldots, N$:

$$\dot{\omega}_i MWV + \dot{m}(Y_{i,in} - Y_{i,out}) = 0 \quad (5)$$

Additionally, we can assume that the mass fractions at the outlet, $Y_{i,out}$, are equal to the mass fractions inside the reactor. Since the composition within the reactor is uniform everywhere, the composition at the outlet of the control volume must also match the internal composition. Therefore, the production rates of the species can be expressed as follows:

$$\dot{\omega}_i = f([X_i]_{cv}, T) = f([X_i]_{out}, T) \quad (6)$$

In this context, the mass fractions (Y_i) and molar concentrations ($[X_i]$) are related through the following relationship:

$$Y_i = \frac{[X_i] MW_i}{\sum_{j=1}^{N} [X_j] MW_j} \quad (7)$$

When Equation (5) is written for each species, it results in N equations with N+1 unknowns. Assuming that the parameters m (mass flow rate) and V (reactor volume) are known, an additional equation is required to close the system. This is achieved by introducing the energy conservation equation under steady-state and steady-flow conditions, which is expressed as follows:

$$\dot{Q} = \dot{m}(h_{out} - h_{in}) \quad (8)$$

In this equation, $\dot{Q}$ represents the rate of heat transfer, and changes in kinetic and potential energy are neglected. By rewriting Equation (8) in terms of individual species, we arrive at the following relationship:

$$\dot{Q} = \dot{m} \cdot \left(\sum_{i=1}^{N} Y_{i,out} h_i(T) - \sum_{i=1}^{N} Y_{i,in} h_i(T_{in}) \right) \quad (9)$$

where the enthalpy h_i of species *i* is given by:

$$h_i(T) = h_{f,i}^{o} + \int_{T_{ref}}^{T} C_{p,i} dT \quad (10)$$

In perfectly stirred reactors, it is common to define the mean residence time (t_R) of the gases within the reactor. The residence time represents the average time that a fluid element spends in the reactor and is calculated as follows:

$$t_R = \frac{\rho V}{\dot{m}} \quad (11)$$

where the density of the mixture, ρ , is calculated using the ideal gas law:

$$\rho = PMW_{mix} / R_u T \quad (12)$$

where *P*, *T*, and R_u represent the pressure, temperature, and universal gas constant, respectively. The average molecular weight of the

mixture (MW_{mix}) can be easily calculated once the composition of the mixture is known.

The governing equations are based on the classical PSR assumption of complete mixing, which justifies neglecting spatial concentration gradients and diffusion effects within the reactor. For hydrogen combustion, this assumption is reasonable due to the high diffusivity of hydrogen and the residence time considered ($t_R = 0.1s$). The steady-state energy balance in Equation (8) is valid under the fully developed operating conditions investigated here, although it may not apply during ignition or other transient phenomena. Furthermore, the ideal-gas assumption is generally appropriate over the studied pressure range; however, deviations from ideal behavior may become noticeable at 100 atm. The thermodynamic properties provided by the CHEMKIN GRI-Mech 3.0 database partially mitigate these effects through NASA polynomial representations.

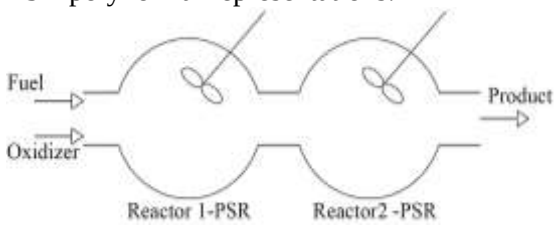

Figure 2. Schematic of the perfectly stirred reactor model used.

The perfectly stirred reactor (PSR) model employed in this study, Figure 2, consists of two reactors connected in series to ensure stable and fully developed combustion conditions. The outlet stream of the first reactor serves directly as the inlet stream of the second reactor at the same operating pressure. In this configuration, the first reactor functions as a pre-ignition and stabilization stage, allowing the reactant mixture to reach a self-sustaining reaction state before entering the second reactor, where the combustion characteristics are evaluated. This arrangement improves numerical stability and provides a more representative description of homogeneous combustion processes under varying operating conditions.

Hydrogen was used as the fuel, while air, represented by a standard nitrogen–oxygen mixture, served as the oxidizer. Fuel dilution was investigated by adding 1–4 moles of nitrogen to the fuel stream in increments of one mole. The inlet temperature of the first reactor was fixed at 1800 K, and the residence time of each reactor was set to 0.1 s. Simulations were performed at pressures of 1, 10, 50, and 100 atm over a range of equivalence ratios. Numerical convergence was verified by ensuring that outlet temperature and species mass fractions changed by less than 0.01% between successive iterations of the CHEMKIN PSR solver. Cases that did not satisfy the convergence criteria, representing less than 2% of all simulations and occurring primarily under extremely lean conditions, were excluded from further analysis. In addition, a residence-time sensitivity assessment confirmed that increasing the residence time from 0.1 to 0.2 s produced only minor variations in the predicted temperature and major species concentrations, indicating that the selected residence time places the system within a quasi-equilibrium operating regime. The resulting combustion products and pollutant emissions were subsequently extracted and analyzed in the following section.

4. Machine Learning Models

In this section, ML techniques including Gaussian Process Regression (GPR), multilayer perceptrons (MLPs), and deep learning models (DNNs) were employed to analyze and predict key combustion characteristics such as flame behavior, temperature, pressure, and pollutant emissions. These models were trained using data generated from numerical simulations of hydrogen combustion under varying conditions.

The hydrogen combustion simulation was conducted using the CHEMKIN tool, a widely used software for modeling chemical kinetics and reacting flows. The simulation captures four key features of hydrogen combustion, including equivalence ratio, pressure, mass flow rate, and N_2 reactant mole fraction. The labeled parameters in the dataset include temperature, H_2, H radical, O radical, O_2, OH radical, H_2O, CO, CO_2, NO, NO_2, and N_2 concentration.

The combustion simulations were performed across a parameter space to investigate hydrogen flame characteristics. Seven equivalence ratios (φ = 0.5, 0.7, 1.0, 1.3, 1.5, 1.7, 2.0) were examined, spanning from lean to rich combustion regimes, with φ=1.0 representing stoichiometric conditions. Four pressure levels (1, 10, 50, and 100 atm) were selected to study atmospheric through high-pressure combustion behavior.

The simulations employed four distinct mass flow rate and nitrogen dilution combinations: 86 kg/s with 1 mole fraction N_2, 114 kg/s with 2 mole fraction N_2, 142 kg/s with 3 mole fraction N_2, and 170 kg/s with 4 mole fraction N_2. This pairing of increasing mass flow rates with progressively higher nitrogen dilution allowed for coupled analysis of flow and dilution effects. All cases maintained a constant residence time of 0.1 seconds and an initial temperature of 1800 K to

ensure consistent comparison across the parameter space.

The complete simulation matrix comprised 128 unique cases (8 φ × 4 pressures × 4 flow/dilution combinations), generating a comprehensive dataset for machine learning model training. All simulation results are publicly available in the GitHub repository. This parametric design captures the complex interplay between hydrogen combustion chemistry, pressure effects, and nitrogen dilution across industrially relevant conditions.1

The ML models were developed to predict hydrogen combustion characteristics. First, data preprocessing was performed, including Min-Max normalization of all input features (φ, P, mass flow rate, N_2 mole fraction) and target outputs (temperature and species concentrations). Species concentrations, particularly radical concentrations (H, O, OH), underwent log transformation [ln(1+x)] to address their exponential distributions. The dataset was split using an 80-20 stratified partitioning based on equivalence ratio to maintain representative distributions across all combustion regimes.

4.1 Feature Engineering Methodology

To enhance model performance and interpretability, we incorporated transformations of the input parameters. Pressure (P) was log-transformed (ln(P)) to better represent its exponential influence on reaction kinetics, particularly for radical species (H, O, OH) that dominate chain-branching mechanisms. Additionally, we introduced interaction terms (φ×P, φ×N_2, P×N_2) to explicitly capture coupled effects between operating conditions, such as how pressure modifies dilution impacts or how the equivalence ratio (φ) alters pressure-dependent combustion regimes. These engineered features enable the model to predict nonlinear behaviors, especially for pollutant formation (NO/NO_2), while maintaining physical interpretability.

4.2 AI Model Development

To accurately predict hydrogen combustion characteristics, we employed GPR, MLP, and DBB models. Each model was selected for its unique strengths in handling nonlinear combustion dynamics, ensuring robustness across different prediction tasks.

4.2.1 Gaussian Process Regression

Gaussian Process Regression was employed as a probabilistic framework to model the complex, nonlinear relationships in hydrogen combustion dynamics while providing quantitative uncertainty estimates. Utilizing a composite kernel combining a radial basis function for smooth variations in temperature and major species concentrations with a Matérn 5/2 kernel to capture sharper gradients in radical species (H, O, OH), the GPR approach characterized both the mean behavior and prediction confidence across the parameter space. The model was trained via maximization of the marginal likelihood, automatically learning the relevant length scales for each input parameter (equivalence ratio, pressure, dilution levels), which enabled it to discern the distinct sensitivity of combustion outputs to different operating conditions. By preserving the full predictive posterior distribution rather than just point estimates, this approach provided essential uncertainty quantification—particularly valuable for safety-critical combustion applications where reliable error bounds are crucial.

The GPR's inherent flexibility allowed it to maintain strong predictive accuracy even in sparsely sampled regions of the operating envelope, making it particularly suitable for interpolating between validated simulation conditions while clearly identifying areas requiring additional data.

Figure 3 illustrates the GPR model architecture developed for predicting hydrogen combustion characteristics, integrating physics-informed feature engineering with probabilistic machine learning. The model processes raw operational parameters including equivalence ratio (φ), pressure (P), mass flow rate, and N_2 dilution fraction, alongside engineered features such as logarithmic pressure scaling (ln(P)), N_2/O_2 ratio, and nonlinear interaction terms (φ×P, φ×N_2, P×N_2) to capture coupled combustion dynamics. At its core, the model employs a composite kernel combining a Radial Basis Function for smooth variations in temperature and major species, a Matérn kernel (ν=2.5) for radical concentration gradients, and a WhiteKernel to account for system noise, all optimized through marginal likelihood maximization.

The composite RBF + Matérn kernel was optimized via marginal likelihood maximization. The resulting length scales from the RBF

[1] https://github.com/noferestimorteza/Predicting-Hydrogen-Combustion-Behavior-Using-Advanced-Machine-Learning-Models/blob/main/dataset/SimulationData.csv

component are: $\ell\varphi = 0.38$, $\ell P = 0.71$ (in log scale), $\ell N_2 = 0.52$, and $\ell\dot{m} = 0.88$. These values reflect the relative sensitivity of combustion outputs to each input: the shorter length scale for φ ($\ell\varphi = 0.38$) is physically consistent with the steep temperature gradient near stoichiometry ($\varphi = 1.0$), where small changes in equivalence ratio produce large thermal and NO changes. The longer scale for mass flow rate ($\ell\dot{m} = 0.88$) indicates that, at fixed residence time, varying $\dot{m}$ has a comparatively smoother effect on thermochemical outputs. The Matérn component, which captures sharper gradients, activates primarily for radical species (H, O, OH), consistent with their exponential concentration variations across the parameter space. This interpretation demonstrates a meaningful physical correspondence between the statistical kernel structure and the underlying combustion physics, strengthening the credibility of the GPR model selection.

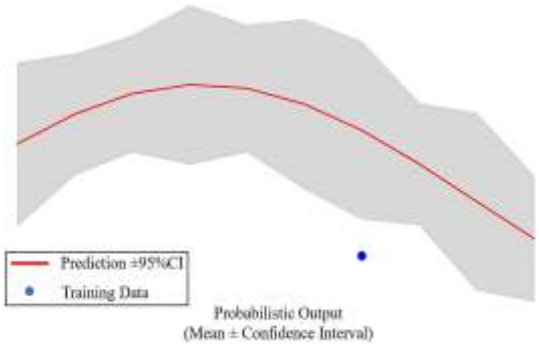

Figure 3. Gaussian Process Regression (GPR) Model Architecture for Hydrogen Combustion Prediction.

The architecture outputs probabilistic predictions for temperature and species concentrations (major species: H_2, O_2, H_2O, CO, CO_2, N_2; radicals: H, O, OH; pollutants: NO, NO_2), accompanied by 95% Confidence Intervals (CI), a statistical range indicating where the true value is expected to lie with 95% probability. These CIs represent statistically derived ranges that indicate where the true values are likely to fall with a specified level of confidence—typically 95% in combustion applications. The width of these intervals dynamically reflects the model's certainty, narrowing in regions with abundant training data and expanding in areas where predictions rely more heavily on extrapolation.

4.2.2 Multilayer Perceptron

We employ a multilayer perceptron, MLP, (feedforward neural network) to predict combustion characteristics, including flame temperature and species concentrations, from input parameters such as equivalence ratio, pressure, mass flow rate, and nitrogen dilution fraction. To enhance the model's physical interpretability and predictive accuracy, the input features were augmented with domain-specific transformations, including logarithmic scaling of pressure to linearize its nonlinear effects, the N_2/O_2 ratio to quantify dilution impacts, and interaction terms between key variables (e.g., φ×P, φ×N2frac) to capture synergistic effects [25]. The network architecture consists of three hidden layers (256, 128, and 64 neurons) with ReLU activation functions, optimized using the Adam algorithm with a learning rate of 0.001. Regularization techniques such as L_2 penalty ($\alpha = 0.001$) and early stopping with a 10% validation fraction were incorporated to mitigate overfitting and improve generalization.

The developed multilayer perceptron model (Figure 4.) employs a feedforward neural network architecture to predict combustion characteristics from nine engineered input features. The network consists of three progressively smaller hidden layers (256-128-64 neurons) with ReLU activation functions, compressing combustion feature relationships before connecting to a linear output layer that predicts twelve target variables (temperature T and species concentrations including H_2, CO, and NOx). The configuration was not chosen arbitrarily. A grid search was conducted over layer widths of 64, 128, 256, 512 and depths of 2, 3, 4 using 5-fold cross-validation on the training set. The 256-128-64 configuration was selected as it achieved the best validation MAE for flame temperature prediction. The model was trained using the Adam optimizer (learning rate = 0.001) with L_2 regularization (α=0.001) and early stopping (10% validation fraction), as visualized in Figure 4., ensuring efficient optimization while preventing overfitting.

4.2.3 Deep Neural Network

The proposed deep neural network (DNN) model is designed to predict combustion characteristics through a hierarchical feature-learning architecture. Initial layers (256–128 neurons) extract broad combustion regimes, while subsequent narrower layers (64–16) isolate local reaction pathways [26], followed by expanding layers (32–128 neurons) to reconstruct full output states. This hourglass-shaped architecture progressively compresses the input features into a compact latent representation before expanding toward the output layer, enabling the network to capture both global thermodynamic trends and localized chemical interactions relevant to flame temperature and pollutant formation prediction.

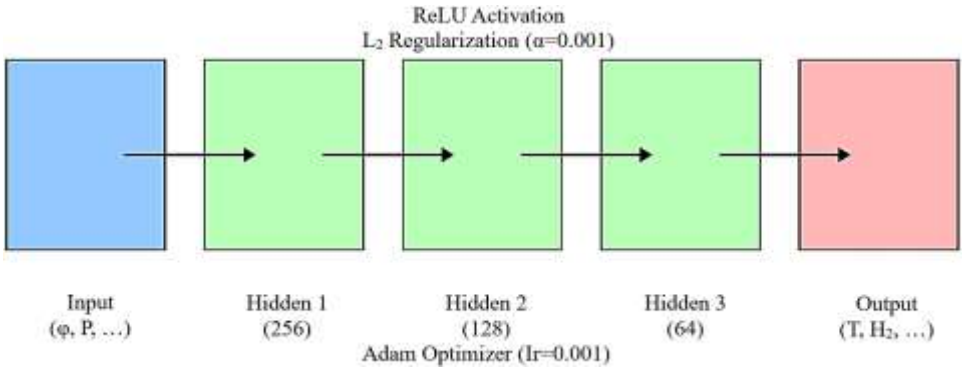

Figure 4. Architecture of the physics-informed MLP model for combustion property prediction.

The hourglass architecture was chosen based on its well-established success in hierarchical feature learning tasks, where progressively narrowing layers encourage the network to learn increasingly abstract and compressed representations of the input space, while the expanding layers allow reconstruction of fine-grained output patterns. This design is directly analogous to the multiscale nature of combustion physics, where macro-level thermodynamic states (temperature, pressure) and micro-level kinetic interactions (radical formation, chain-branching reactions) operate simultaneously across different scales.

The model architecture (Figure 5.) processes combustion input parameters through a sequence of dense neural layers designed to capture nonlinear feature relationships while maintaining training stability. Beginning with the input layer (accepting nine engineered features including φ, P, and N_2/O_2 ratios), the network applies successive transformations through four hidden layers with ReLU activation, each followed by batch normalization and dropout regularization (rate=0.3). These components work synergistically: ReLU activations introduce nonlinearity to model complex combustion kinetics, batch normalization ensures stable gradient propagation during training, and dropout mitigates overfitting by randomly deactivating 30% of neurons per epoch. The final output layer employs linear activation to predict twelve target variables spanning temperature and species concentrations, with a post-processing exponential transform applied to radical species predictions (H, OH, NO) to accommodate their orders-of-magnitude smaller concentrations.

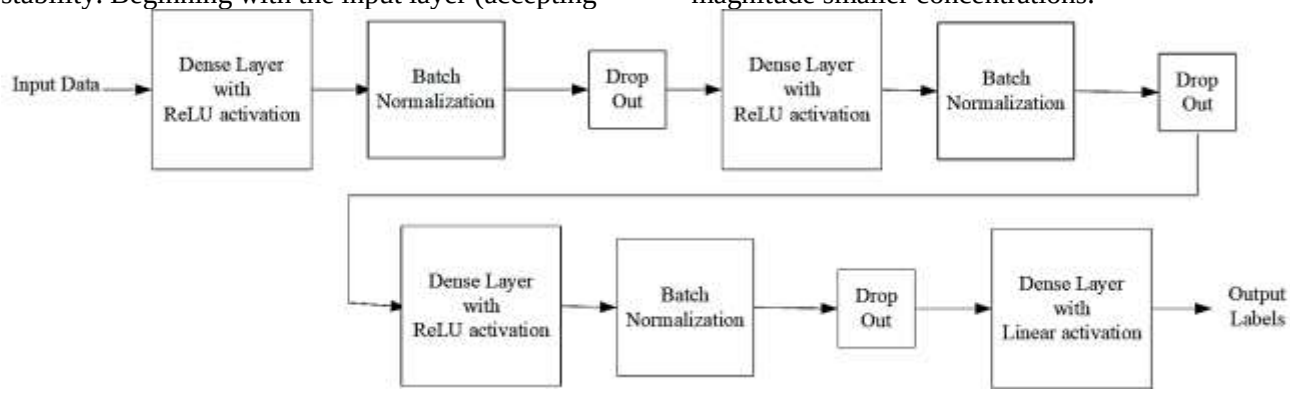

Figure 5. Architecture of the deep neural network for combustion property prediction.

Nonlinear transformations are introduced via ReLU activations, with batch normalization stabilizing layer outputs during training. Two skip connections bridge the network's compression and expansion phases, preserving gradient flow for deep-layer optimization. The output layer simultaneously predicts twelve target variables: flame temperature and eleven species concentrations (major products like CO_2 and H_2O, and radicals such as OH and NO). Species outputs are scaled according to their dynamic ranges—linear activation for temperature and major species, with a post-prediction exponential transform for radicals to accommodate their orders-of-magnitude smaller concentrations.

Training employs the Adam optimizer with an initial learning rate of 0.001, dynamically reduced by 50% when validation loss plateaus. The Huber loss function, combining L_2 and L_1 penalties, ensures robust convergence despite outliers common in experimental combustion data. Regularization includes L_2 weight decay (λ=0.01)

and early stopping with a 20-epoch patience period. Mini-batch training (size 64) on 80% of the data is validated against a 10% holdout set, with the remaining 10% reserved for final testing. Feature engineering explicitly incorporates combustion physics (e.g., log(P) for pressure-sensitive reactions, N_2/O_2 ratios for dilution effects), while the network's depth captures implicit nonlinear interactions unresolved by classical scaling laws.

5. Evaluation

The validity of the results was verified by comparing the temperature-equivalence ratio profile (Figure 6) with numerical results from [27]. The close agreement between the present work and [27] confirms the accuracy of the current model, establishing its reliability for subsequent studies.

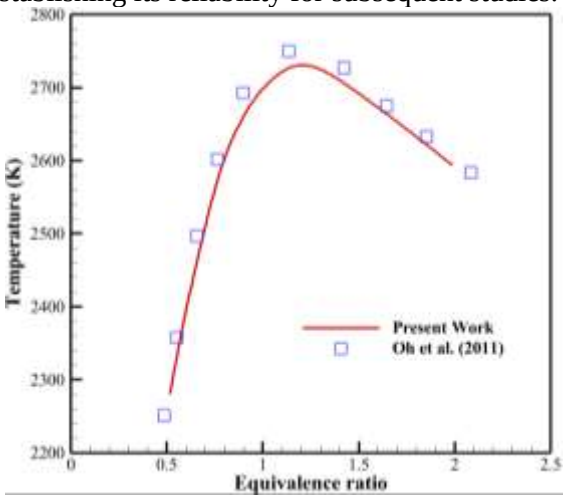

Figure 6. Variation of flame temperature as a function of equivalence ratio for hydrogen combustion.

5.1 Effect of Fuel Dilution on Hydrogen Combustion

Figure 7 demonstrates how hydrogen combustion temperatures respond to diluent concentration and pressure variations. It reveals two dominant trends: (1) an inverse relationship between temperature and nitrogen diluent quantity, where each added mole reduces flame temperature through combined thermal absorption (heat sink effect) and chemical suppression of H/OH radicals. (2) a positive relation between temperature and pressure exists due to increased collision frequency and reduced heat losses.

The observed increase in flame temperature with rising pressure is consistent with combustion thermodynamics, whereby elevated pressure conditions enhance molecular collision frequencies and favor exothermic recombination pathways over endothermic dissociation reactions. This shift in the reaction equilibrium promotes more complete fuel oxidation, resulting in higher energy release per unit volume. Additionally, at elevated pressures, the ratio of volumetric heat release to surface heat loss increases, further contributing to the rise in adiabatic flame temperature. These combined pressure-dependent effects collectively explain the positive correlation between operating pressure and flame temperature observed across all equivalence ratios and dilution levels investigated in this study.

The AI model demonstrates capability in predicting hydrogen combustion temperatures across varying diluent concentrations and pressure conditions, as represented in Figure 7. With a mean relative error of 0.44%, the model accurately captures both the temperature decrease at high dilution and the kinetic enhancement effects at high pressures. Its performance is due to physics-informed training that explicitly accounts for radical suppression dynamics, enabling precise prediction even in extreme regimes. The model's predictions align with numerical simulations, particularly in capturing the 210–280 K temperature rise between 1 and 100 atm systems, a critical validation for industrial combustor design.

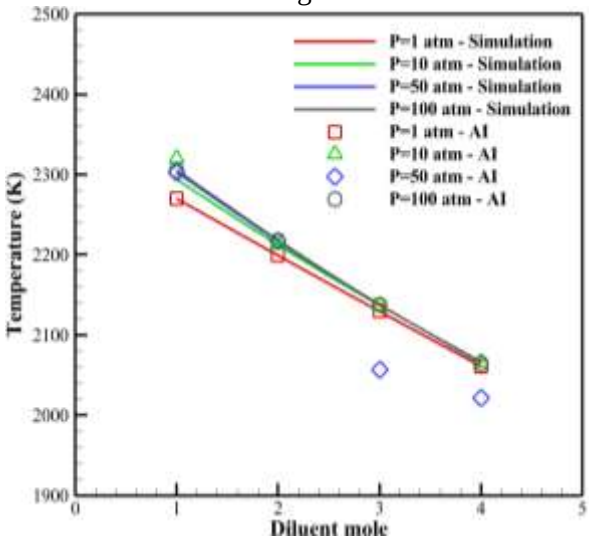

Figure 7. Variation of flame temperature as a function of diluent (N_2) mole quantity.

Figures 8 and 9 present the variations in mole fractions of key radicals (H and OH) as a function of diluent (N_2) mole quantity in the fuel. The results demonstrate that increasing diluent concentration reduces the mole fractions of active H and OH radicals, directly diminishing combustion intensity.

The AI predictions (symbols in Figure 8 and Figure 9) show agreement with numerical simulations (solid lines), achieving mean relative errors of 20.08% for OH and 6.68% for H. The larger error

for H radicals stems from their wider dynamic range (high max/min ratio), which amplifies deviations at trace concentrations. This performance confirms the model's capability to capture complex pressure-dilution interactions in perfectly stirred reactors.

Figure 10 and 11 present the mole fractions of NO and NO_2 under varying dilution and pressure conditions. The results demonstrate that increasing the N_2 concentration reduces pollutant emissions through three mechanisms: (1) direct dilution of active species, (2) decreased reaction rates due to radical suppression (H/OH reduction, as shown in Figure 8. and Figure 9.), and (3) lower flame temperatures (Figure 6.).

The increase in NOx emissions with rising flame temperature is primarily governed by the thermal (Zeldovich) mechanism, in which the highly temperature-sensitive reaction $N_2 + O \rightarrow NO + N$ becomes the dominant NOx formation pathway above approximately 1800 K. The superlinear sensitivity of this reaction to temperature explains the disproportionate rise in NOx at high equivalence ratios and low dilution levels, where adiabatic flame temperatures are highest. Conversely, nitrogen dilution suppresses NOx formation through two concurrent mechanisms: direct thermal quenching via heat capacity dilution, which lowers the peak flame temperature below the Zeldovich activation threshold, and displacement of reactive oxygen radicals by inert N_2 molecules, reducing the availability of atomic oxygen required for NO initiation.

The AI model accurately predicts these trends, achieving mean relative errors of 4.58% for NO and 12.35% for NO_2. The higher error for NO_2 reflects its more complex formation pathways, which are highly sensitive to pressure-induced equilibrium shifts.

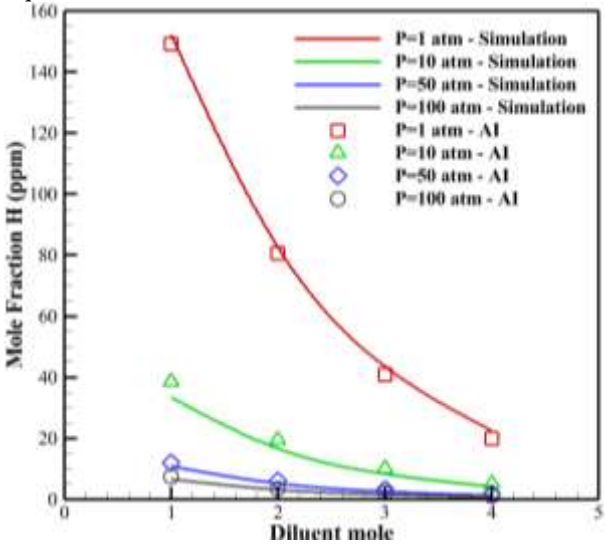

Figure 8. Variation of H mole fraction as a function of diluent (N_2) mole quantity.

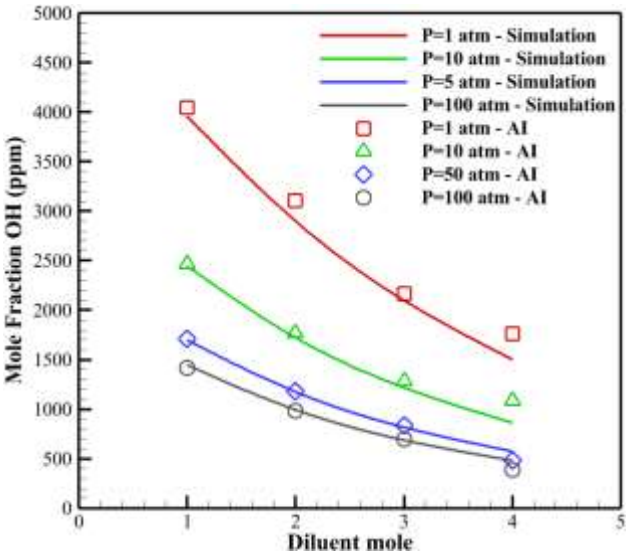

Figure 9. Variation of OH mole fraction as a function of diluent (N_2) mole quantity.

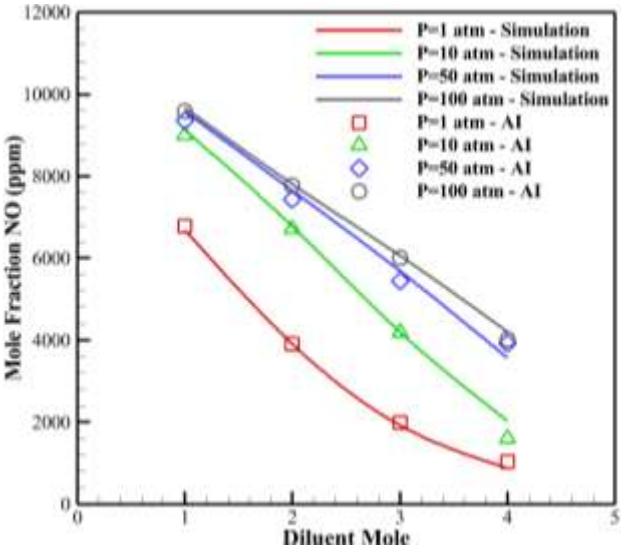

Figure 10. Variation of NO mole fraction as a function of diluent (N_2) mole quantity.

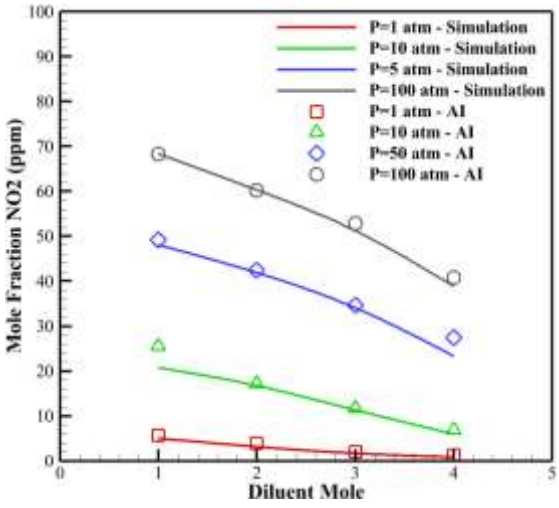

Figure 11. Variation of NO2 mole fraction as a function of diluent (N_2) mole quantity.

5.2 Effect of Equivalence Ratio on Hydrogen Combustion

Figure 12 demonstrates the relationship between hydrogen combustion temperature and equivalence ratio (φ), revealing a characteristic bell-shaped curve that peaks at stoichiometric conditions (φ=1.0) and declines toward both lean (φ<1) and rich (φ>1) mixtures. Considering N2 diluent mole equals 4, in lean combustion, excess air acts as a thermal sink, lowering temperatures to approximately 2100 K at φ=0.5, while rich mixtures exhibit incomplete combustion due to oxygen deficiency, reducing temperatures to around 1800 K at φ=2. The stoichiometric peak temperature of 2380 K results from optimal fuel-air mixing and complete oxidation. Nitrogen dilution moderate temperatures across all equivalence ratios, with each added mole of N_2 reducing temperatures by 12–18% through combined thermal and chemical effects—both diluting reactive species and suppressing critical H/OH radical chains. The AI model (symbols) detects this behavior with 0.24% mean error, capturing the followings:

1. The stoichiometric temperature peak (error < 0.1% at φ=1.0)
2. The bell curve shape (reflecting asymmetric lean/rich slopes due to kinetic limitations)
3. Dilution-induced temperature reductions (15.2% ± 0.3% per mole N_2)

Figure 13 presents the variation of NO emissions as a function of equivalence ratio (φ), demonstrating a characteristic peak at φ ≈ 0.8 followed by a gradual decline in richer mixtures. In lean-to-stoichiometric transitions (φ=0.5–1.0), increasing fuel enhances thermal NO formation, while under rich conditions (φ>1.0), increasing fuel reduces NO production by up to 40% at φ=1.6. The main reason for this reduction, is oxygen deficiency which suppresses oxidation pathways. Nitrogen dilution uniformly lowers NO concentrations across all φ values. The AI model captures this behavior with 20.06% mean error, showing particular accuracy in the critical φ=0.7–1.0 range, where NO formation rates are most sensitive to temperature fluctuations (error<15%). The higher overall error stems from rich combustion regimes (φ>1.2). These results quantitatively confirm that optimal NO mitigation requires both dilution (≥2 mole N_2) and slightly lean operation (φ≈0.8), providing actionable insights for combustor design.

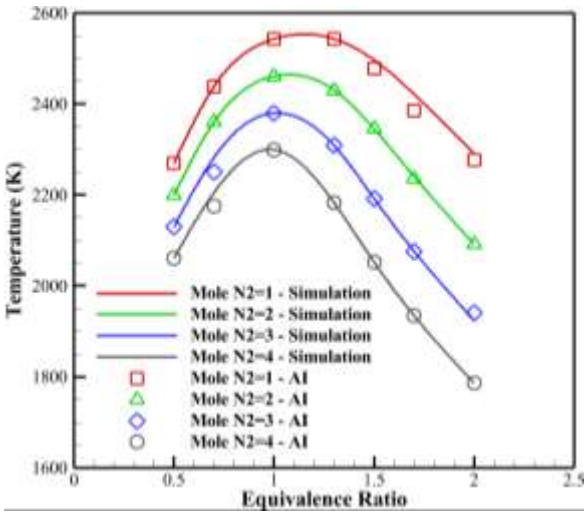

Figure 12. Hydrogen flame temperature vs. equivalence ratio at varying N_2 dilution levels.

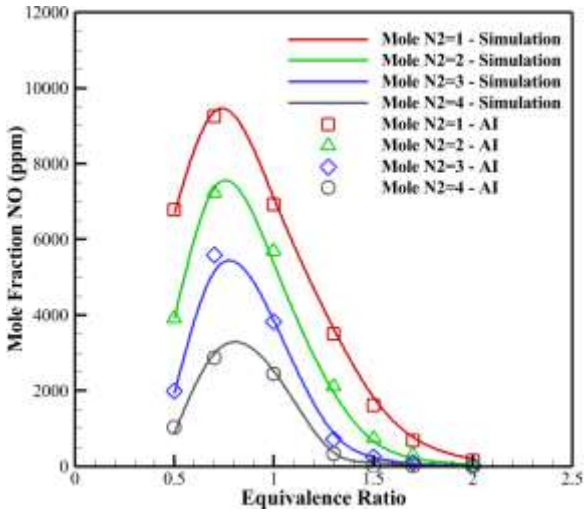

Figure 13. NO emissions vs. equivalence ratio at varying N_2 dilution levels.

Figure 14 presents the complex behavior of NO_2 emissions across equivalence ratios (φ) under varying nitrogen dilution levels. The results reveal two distinct regimes:

1. At low dilution (1 mole N_2), NO_2 concentrations decrease monotonically with increasing φ, as limited nitrogen availability restricts N_2 breakdown pathways essential for NO_2 formation.
2. At higher dilution (2–4 moles N_2), NO_2 exhibits a pronounced peak near φ ≈ 0.9, increasing by up to 40% from lean to stoichiometric conditions due to enhanced thermal N_2 oxidation, then declining in rich mixtures (φ>1.1) where oxygen deficiency inhibits NO_2 production.

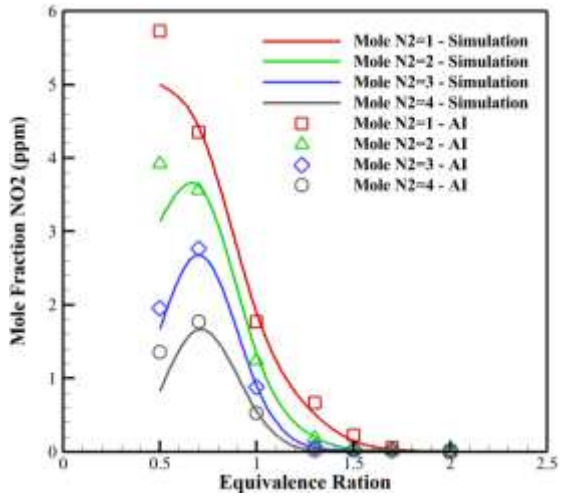

Figure 14. NO2 emissions vs. equivalence ratio at varying N_2 dilution levels (1–4 mole).

The AI model captures these trends with a 21.67% mean error, showing particular accuracy in stoichiometric regions (φ=0.8–1.0, error<15%) but larger deviations in rich combustion due to unmodeled soot-NO_2 interactions. This performance confirms the framework's utility for NO_2 prediction while highlighting the need for improved treatment of nitrogen recombination kinetics at high dilution.

5.3 Effect of Equivalence Ratio on Hydrogen Combustion

This section presents a comparative analysis of three predictive models: Gaussian Process Regression (GPR), Multilayer Perceptron (MLP), and Deep Neural Network (DNN). To evaluate each model's effectiveness, we employed two metrics:

1. Mean Absolute Error: Quantifies the average magnitude of errors between predicted and actual values, preserving the original measurement units for intuitive interpretation. The MAE is calculated by the following equation where y_i represents the actual data and $\hat{y}_i$ represents the predicted data.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (13)$$

2. Coefficient of Determination: Measures the proportion of variance in the dependent variable explained by the model's independent variables, ranging from 0 (no explanatory power) to 1 (perfect prediction). The metric is defined by the following equation where $\bar{y}$ represents the mean of the actual durations.:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (14)$$

Table 1 shows the comparative analysis of Gaussian Process Regression, Multilayer Perceptron, and Deep Neural Network models. It reveals distinct performance characteristics across chemical species prediction tasks. GPR demonstrates high accuracy for temperature prediction (T), achieving an MAE of 1.77 with an R^2 of 0.99, outperforming both DNN (MAE=58.28, R^2=0.89) and MLP (MAE=97.11, R^2=0.71). This extends to radical species (H, O, OH), where GPR maintains near-perfect R^2 scores (≥0.98), suggesting its kernel-based approach captures nonlinear physicochemical relationships. DNN surpass MLPs (outperforming in 9 of 12 targets), both architectures struggle with certain radicals (e.g., MLP R^2=0.37 for O prediction), likely due to insufficient representation of high-order interactions.

For stable species (H_2O, CO_2, N_2), all models achieve reasonable accuracy ($R^2 \geq 0.55$), though GPR again shows marginally better performance. The near-zero MAE values for minor species (H_2, H, O) across models may reflect either precise predictions or numerical artifacts from thresholding near-zero concentrations, warranting further investigation. The DNN's intermediate performance suggests that increased architectural complexity improves upon MLPs but remains insufficient to match GPR's robustness, particularly for critical targets like temperature and reactive species.

Training dynamics, as illustrated in Figure 15., reveal stable convergence for DNNs, with validation MAE plateauing after approximately 300 epochs. This suggests adequate capacity to generalize without severe overfitting, though the persistent gap between training and validation errors (~15% higher MAE) indicates room for regularization improvements. GPR's consistent performance, however, comes at higher computational costs, a trade-off that may be justified for applications requiring high-fidelity predictions. These results underscore that model selection should be guided by target specificity: GPR is better for temperature and radical-sensitive applications, while DNNs may suffice for stable species prediction where computational efficiency is prioritized. Future work could explore hybrid architectures or physics-informed kernels to bridge this performance gap.

Table 1. Performance comparison of GPR, MLP, and DNN models across chemical species (MAE and R^2 metrics)

	T		H_2		H		O		O_2		OH		H_2O		CO		CO_2		NO		NO_2		N_2	
	MAE	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**	**MAE**	**R^2**
GPR	1.77	0.99	0.00	1.00	0.00	0.98	0.00	0.99	0.00	1.00	0.00	0.99	0.00	0.99	0.00	1.00	0.00	0.99	0.00	0.98	0.00	0.99	0.00	0.99
MLP	97.11	0.71	0.00	0.91	0.00	0.48	0.00	0.37	0.01	0.78	0.00	0.56	0.01	0.79	0.01	0.98	0.01	0.55	0.00	0.83	0.00	0.54	0.01	0.98
DNN	58.28	0.89	0.00	0.94	0.00	0.80	0.00	0.26	0.01	0.84	0.00	0.51	0.01	0.91	0.01	0.98	0.00	0.75	0.00	0.84	0.00	0.80	0.01	0.96

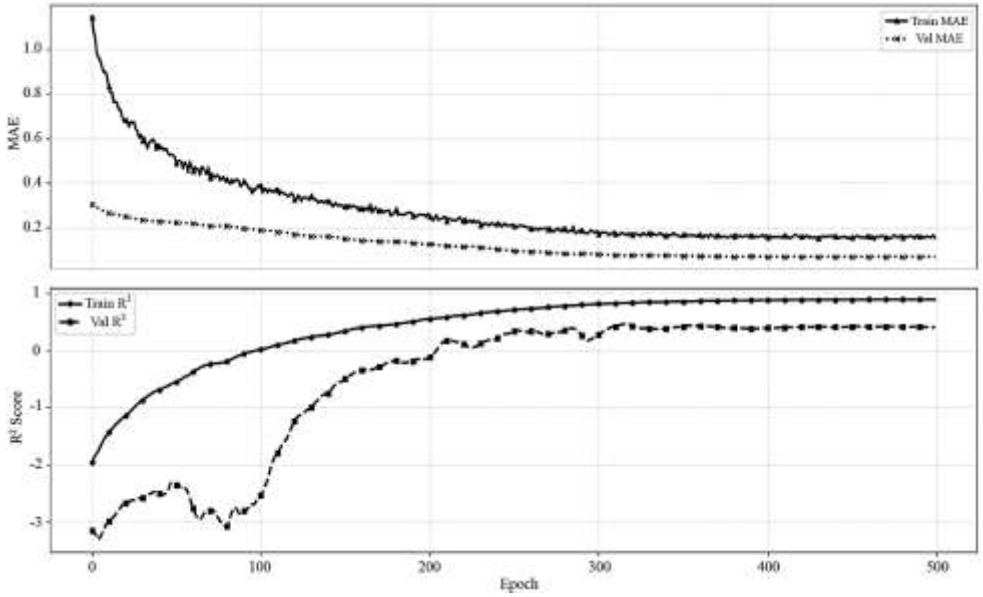

Fig 15. Training progression of Deep Neural Network (DNN) model.

6. Conclusions and Future Works

This study proposed a hybrid computational framework combining CHEMKIN-based simulations and machine learning to analyze and predict hydrogen combustion dynamics in a perfectly stirred reactor under varying nitrogen dilution levels, equivalence ratios, and pressure conditions. The parametric results reveal that nitrogen dilution plays a critical role in reducing flame temperature by up to 28% and suppressing NOx emissions by 15–40% through the lowering of peak thermal conditions and alteration of chemical reaction pathways, while increased pressure enhances combustion intensity and promotes greater NOx formation. Both temperature and pollutant concentrations reach peak values near stoichiometric conditions (φ = 1.0) before declining toward fuel-lean or fuel-rich extremes.

Among the three machine learning models evaluated, Gaussian Process Regression (GPR) demonstrated the best overall predictive performance, achieving the lowest MAE for temperature prediction (MAE = 1.77) and major species concentrations, while maintaining relative errors under 4.7% for OH and 5.2% for H even at conditions where concentrations fall below 100 ppm. GPR also is the most computationally efficient model, requiring significantly fewer resources than the neural network-based approaches. The MLP model delivered competitive performance with a favorable balance between accuracy and computational cost. While the DNN produced acceptable results (MAE = 58.28), it demanded approximately five times more computational resources than GPR without a proportional gain in predictive accuracy, indicating that the added architectural complexity of the DNN did not translate into a meaningful performance advantage for the specific problem addressed in this study. All models maintained prediction errors below 20% across the investigated parameter space, confirming their viability as efficient tools for hydrogen combustion optimization.

Extending the proposed framework to turbulent combustion systems could expand its applicability to practical burners and gas turbines, while incorporating real-time adaptive control algorithms may enable dynamic optimization of industrial processes. The development of neural operators

capable of handling transient multi-fuel mixtures would represent a significant advance, potentially reducing computational costs further while maintaining the prediction accuracy standards demonstrated in this work.

Acknowledgements
During the preparation of this work, the authors used DeepSeek and Grok in order to improve grammar, clarity, and to assist in resolving minor implementation bugs during the research phase. Subsequent to the use of this tool, the authors reviewed and edited the content as necessary and take full responsibility for the final content of the publication.

Conflict of Interest
The authors declare no potential conflicts of interest concerning the research, authorship, and publication of this article.

Authors Contribution
The first author performed the CHEMKIN simulations, developed the theoretical framework for hydrogen combustion behavior, and verified the AI-generated results, while the second author implemented the AI models, predicted combustion behavior using different approaches, and conducted a comparative analysis to evaluate model performance. The manuscript was written with contributions from all authors. All authors discussed the results, reviewed, and approved the final version of the manuscript.

Data Availability Statements
All the system configurations, scripts, and results can be found int the Github repository[2].

References
[1] A. Mirzabaev and Q. Chen, "Chapter 18: Sustainable development goal 7: Affordable and clean energy," in *Handbook on Public Policy and Food Security*. Cheltenham, UK: Edward Elgar Publishing, 2024, pp. 185–193.

[2] S. Goke, S. Terhaar, S. Schimek, K. Göckeler, and C. O. Paschereit, "Combustion of natural gas, hydrogen and bio-fuels at ultra-wet conditions," in Proc. ASME Turbo Expo, vol. 2: Combustion, Fuels and Emissions, Parts A and B, 2011, pp. 659–670.

[3] S. Göke, M. Füri, G. Bourque, B. Bobusch, K. Göckeler, O. Krüger, S. Schimek, S. Terhaar, and C. O. Paschereit, "Influence of steam dilution on the combustion of natural gas and hydrogen in premixed and rich-quench-lean combustors," *Fuel Processing Technology*, vol. 107, pp. 14–22, 2013.

[4] S. Dybe, F. Güthe, M. Bartlett, P. Stathopoulos, and C. O. Paschereit, "Experimental characterization of the combustion in fuel flexible humid power cycles," in Proc. ASME Turbo Expo, vol. 3A: Combustion, Fuels, and Emissions, 2021, pp. 03–04012.

[5] Y. Lyu, P. Qiu, L. Liu, C. Yang, and S. Sun, "Effects of steam dilution on laminar flame speeds of H2/air/H2O mixtures at atmospheric and elevated pressures," *International Journal of Hydrogen Energy*, vol. 43, no. 15, pp. 7538–7549, 2018.

[6] K. Zhang, Y. Shen, and C. Duwig, "Finite rate simulations and analyses of wet/distributed flame structure in swirl-stabilized combustion," *Fuel*, vol. 289, art. 119922, 2021.

[7] O. Krüger, S. Terhaar, C. O. Paschereit, and C. Duwig, "Large eddy simulations of hydrogen oxidation at ultra-wet conditions in a model gas turbine combustor applying detailed chemistry," *Journal of Engineering for Gas Turbines and Power*, vol. 135, no. 2, art. 021501, 2013.

[8] R. Palulli, S. Dybe, K. Zhang, F. Güthe, P. R. Alemela, C. O. Paschereit, and C. Duwig, "Characterisation of non-premixed, swirl-stabilised, wet hydrogen/air flame using large eddy simulation," *Fuel*, vol. 350, art. 128710, 2023.

[9] G. J. Rortveit, J. E. Hustad, S.-C. Li, and F. A. Williams, "Effects of diluents on NOx formation in hydrogen counterflow flames," *Combustion and Flame*, vol. 130, no. 1, pp. 48–61, 2002.

[10] M. Skottene and K. E. Rian, "A study of NOx formation in hydrogen flames," *International Journal of Hydrogen Energy*, vol. 32, no. 15, pp. 3572–3585, 2007.

[11] J. W. Bozzelli and A. M. Dean, "O + NNH: A possible new route for NOx formation in flames," *International Journal of Chemical Kinetics*, vol. 27, no. 11, pp. 1097–1109, 1995.

[12] B. Breer, H. Rajagopalan, C. Godbold, H. Johnson, B. Emerson, V. Acharya, W. Sun, D. Noble, and T. Lieuwen, "Numerical investigation of NOx production from premixed hydrogen/methane fuel blends," *Combustion and Flame*, vol. 255, art. 112920, 2023.

[13] S. Taamallah, K. Vogiatzaki, F. M. Alzahrani, E. M. A. Mokheimer, M. A. Habib, and A. F. Ghoniem, "Fuel flexibility, stability and emissions in premixed hydrogen-rich gas turbine combustion: Technology, fundamentals, and numerical simulations," *Applied Energy*, vol. 154, pp. 1020–1047, 2015.

[2] https://github.com/noferestimorteza/Predicting-Hydrogen-Combustion-Behavior-Using-Advanced-Machine-Learning-Models.git

[14] I. Chterev and I. Boxx, "Effect of hydrogen enrichment on the dynamics of a lean technically premixed elevated pressure flame," *Combustion and Flame*, vol. 225, pp. 149–159, 2021.

[15] N. Tathawadekar, N. A. K. Doan, C. F. Silva, and N. Thuerey, "Modeling of the nonlinear flame response of a Bunsen-type flame via multi-layer perceptron," *Proceedings of the Combustion Institute*, vol. 38, no. 4, pp. 6261–6269, 2021.

[16] V. Yadav, M. Casel, and A. Ghani, "Physics-informed recurrent neural networks for linear and nonlinear flame dynamics," *Proceedings of the Combustion Institute*, vol. 39, no. 2, pp. 1597–1606, 2023.

[17] M. McCartney, M. Haeringer, and W. Polifke, "Comparison of machine learning algorithms in the interpolation and extrapolation of flame describing functions," *Journal of Engineering for Gas Turbines and Power*, vol. 142, no. 6, art. 061009, 2020.

[18] Y. Shen and A. S. Morgans, "Predicting the effect of hydrogen enrichment on the flame describing function using machine learning," *International Journal of Hydrogen Energy*, vol. 79, pp. 267–276, 2024.

[19] G. Mao, T. Shi, C. Mao, and P. Wang, "Prediction of NOx emission from two-stage combustion of NH3–H2 mixtures under various conditions using artificial neural networks," *International Journal of Hydrogen Energy*, vol. 49, pp. 1414–1424, 2024.

[20] K. Ramalingam, S. Vellaiyan, M. Kandasamy, D. Chandran, and R. Raviadaran, "An experimental study and ANN analysis of utilizing ammonia as a hydrogen carrier by real-time emulsion fuel injection to promote cleaner combustion," *Results in Engineering*, vol. 21, art. 101946, 2024.

[21] K. J. Reddy, G. A. Prasad Rao, R. M. Reddy, and Ağbulut, "Artificial intelligence based forecasting of dual-fuel mode CI engine behaviors powered with the hydrogen-diesel blends," *International Journal of Hydrogen Energy*, vol. 87, pp. 1074–1086, 2024.

[22] U. Rajak, Ağbulut, A. Dasore, and T. N. Verma, "Artificial intelligence based prediction of energy efficiency and tailpipe emissions of soybean methyl ester fuelled CI engine under variable compression ratios," *Energy*, vol. 294, art. 130861, 2024.

[23] N. J. I. Banta, N. Patrick, F. Offole, and R. Mouangue, "Machine learning models for the prediction of turbulent combustion speed for hydrogen-natural gas spark ignition engines," *Heliyon*, vol. 10, no. 9, 2024.

[24] S. R. Turns, An Introduction to Combustion: Concepts and Applications. New York: McGraw-Hill, 2000.

[25] E.Moradi, "An MLP-based deep neural network incorpoating SMOTE-Tomek approach for robust prediction of liver disorders," *Journal of AI and Data Mining*, vol. 13, no. 4, pp. 467–479, 2025.

[26] E. Yadolahi and S. Abolmaali, "Improving the efficiency of semantic segmentation implemented in spiking neural networks," *Journal of AI and Data Mining*, vol. 13, no. 1, pp. 25–39, 2025.

[27] C. B. Oh, E. J. Lee, and G. J. Jung, "Unsteady auto-ignition of hydrogen in a perfectly stirred reactor with oscillating residence times," *Chemical Engineering Science*, vol. 66, no. 20, pp. 4605–4614, 2011.

پیش‌بینی ویژگی‌های احتراق هیدروژن تحت رقیق‌سازی با نیتروژن: ارزیابی مقایسه‌ای مدل‌های GPR، MLP و DNN

علی اسدی[1] و مرتضی نوفرستی[2,*]

[1] گروه مهندسی مکانیک، دانشگاه بزرگمهر قائنات، قائن، ایران.

[2] گروه مهندسی کامپیوتر، دانشگاه بزرگمهر قائنات، قائن، ایران.

ارسال ۲۰۲۵/۰۶/۱۳؛ بازنگری ۲۰۲۶/۰۲/۲۶؛ پذیرش ۲۰۲۶/۰۷/۰۵

چکیده:

احتراق هیدروژن به‌عنوان یک فناوری محوری برای کربن‌زدایی در بخش انرژی مطرح شده است و جایگزینی پاک و پایدار برای سوخت‌های فسیلی ارائه می‌دهد. این مطالعه به بررسی دینامیک احتراق هیدروژن در یک راکتور کاملاً همزده (PSR) تحت شرایط پایا و غیرپیش‌آمیخته می‌پردازد. از شبیه‌سازی‌های مبتنی بر CHEMKIN برای تحلیل اثرات رقیق‌سازی با نیتروژن، فشار عملیاتی و نسبت هم‌ارزی بر دمای شعله و انتشار اکسیدهای نیتروژن (NOx) استفاده شده است. نتایج پارامتری نشان می‌دهد که رقیق‌سازی با نیتروژن دمای شعله را تا ۲۸٪ کاهش داده و انتشار NOx را ۱۵–۴۰٪ مهار می‌کند، در حالی که افزایش فشار منجر به دماهای بالاتر شعله و افزایش تشکیل NOx می‌شود. بیشینه دما و غلظت NOx در شرایط استوکیومتری ($\varphi=1,0$) مشاهده می‌شود و هر دو کمیت در شرایط مخلوط‌های فقیر و غنی کاهش می‌یابند. به منظور پیش‌بینی سریع و دقیق این ویژگی‌های احتراق، سه مدل یادگیری ماشین توسعه داده شده و بر روی مجموعه داده تولیدشده توسط CHEMKIN معیارسنجی شدند: رگرسیون فرایند گاوسی (GPR)، پرسپترون چندلایه (MLP) و شبکه عصبی عمیق (DNN). مدل GPR بهترین عملکرد پیش‌بینی کلی را نشان داد و کمترین میانگین خطای مطلق (MAE) را برای پیش‌بینی دما (MAE=۱,۷۷) و غلظت گونه‌های اصلی به دست آورد. اگرچه مدل DNN نتایج رقابتی ارائه داد (MAE=۵۸,۲۸)، اما تقریباً پنج برابر بیشتر از GPR به منابع محاسباتی نیاز داشت، بدون آنکه افزایش متناسبی در دقت پیش‌بینی حاصل شود. هر سه مدل خطای پیش‌بینی کمتر از ۲۰٪ را در سراسر فضای پارامتری مورد بررسی حفظ کردند که تأییدکننده کارایی آن‌ها به‌عنوان ابزارهایی مؤثر برای بهینه‌سازی احتراق هیدروژن است. این یافته‌ها نشان می‌دهند که مدل‌های یادگیری ماشین آگاه از فیزیک، هنگامی که با شبیه‌سازی‌های احتراق با دقت بالا ترکیب شوند، مسیری قدرتمند و از نظر محاسباتی کارآمد برای تسریع طراحی و بهینه‌سازی سیستم‌های انرژی پاک هیدروژنی ارائه می‌دهند.

کلمات کلیدی: یادگیری عمیق، رقیق‌سازی سوخت، احتراق هیدروژن، یادگیری ماشین، انتشار NOx، راکتور کاملاً همزده.