International Journal of Hydrogen Energy 234 (2026) 154875

Contents lists available at ScienceDirect

International Journal of Hydrogen Energy

journal homepage: www.elsevier.com/locate/he

Physics-based analytical correlation for NO_x formation in partially premixed swirl-stabilized lean H_2 combustors

Alam Garcidueñas-Correa *, Francesca De Domenico, Arvind Gangoli Rao

Faculty of Aerospace Engineering, TU Delft, Kluyverweg 1, Delft, 2629HS, South-Holland, The Netherlands

ARTICLE INFO

Dataset link: 10.4121/8df3f562-0ded-4308-89a6-9b3b8a5c1a68

Keywords:
Hydrogen
Nitrogen oxides
Analytical correlation
Swirl-stabilized
Partially-premixed

ABSTRACT

Hydrogen combustion is gaining attention for its potential to enable low-emission energy conversion in gas turbines. Since hydrogen is carbon-free, it produces no carbon-based pollutants and primarily forms water vapor. However, due to its higher adiabatic flame temperature with respect to hydrocarbons, hydrogen combustion is more prone to increased nitrogen oxides (NO_x) formation. Accurately predicting NO_x formation remains a major challenge, particularly when scaling from laboratory experiments to industrial applications. While scaling laws are widely used in fluid dynamics, their application to NO_x emissions is challenging due to the complex nature of NO_x formation. This study develops a semi-empirical, physics-based correlation to estimate NO_x emissions in swirl-stabilized, partially premixed lean hydrogen–air burners. The proposed correlation expresses the Emission Index of NO_x ($EINO_x$) as a function of key operating and design parameters, including fuel mass flow rate, pressure, adiabatic flame temperature, equivalence ratio, residence time and swirl number. It builds upon Westenberg's NO_x formation rate equation, incorporating composition-dependent effects via equivalence ratio. The slow formation nature of NO_x is accounted for via the combustor mean residence time. Additionally, the influence of swirl on the mixing process is modeled through a swirl-modified effective equivalence ratio, acknowledging that while partial premixing is a design choice, swirl intensity can either enhance or disrupt the degree of premixing before combustion. The model's parameters were calibrated using experimental data from published literature. By providing a predictive tool for NO_x scaling across different operating conditions, this model supports the development and design of experiments and devices for hydrogen–air combustion.

1. Introduction

The climate crisis, driven by industrialization, has led to rising greenhouse gas concentrations, causing global warming and environmental instability. The growing demand for fossil fuels worsens these effects, particularly in gas-fired power plants, which are major CO_2 emitters. Hydrogen offers a promising alternative for decarbonizing power generation, as its combustion produces mostly water vapor, eliminating direct CO_2 emissions. If produced renewably, hydrogen could contribute to a more sustainable energy future [1]. Additionally, its higher gravimetric energy density allows for reduced fuel consumption compared to fossil fuels.

Despite these advantages, hydrogen combustion poses important challenges, most notably the formation of nitrogen oxides (NO_x) due to the high combustion temperatures [2]. Hydrogen generally has a higher combustion temperature than that hydrocarbons, therefore, if it is poorly mixed with air, NO_x emissions can be much higher. NO_x emissions are a societal concern since they contribute to air pollution, ozone formation, and acid rain, and are strictly regulated by agencies such as International Civil Aviation Organization (ICAO) and the U.S. Environmental Protection Agency (EPA).

Accurate prediction methods are therefore essential for both advancing research and ensuring compliance with emission standards [3]. In response, both industry and academia are intensifying efforts to study hydrogen combustion, aiming to better understand its processes and to develop effective emission control strategies. Laboratory-scale research plays a key role in uncovering the fundamental mechanisms of hydrogen combustion. However, translating these insights into industrial applications remains complex due to the numerous interacting parameters involved [4].

Traditional similarity theory, widely used in fluid mechanics, is limited when applied to reactive flows, as it requires maintaining numerous similarity parameters [4]. Spalding [4] proposed a partial modeling approach, which focuses on the most critical parameters,

* Corresponding author.
E-mail address: A.U.GarciduenasCorrea@tudelft.nl (A. Garcidueñas-Correa).

https://doi.org/10.1016/j.ijhydene.2026.154875
Received 9 October 2025; Received in revised form 26 February 2026; Accepted 2 April 2026
Available online 20 April 2026

Nomenclature

R_{gas}	Specific gas constant (J kg^{-1} K^{-1})
P_{op}	Operating pressure (bar)
P_0	Reference pressure (bar)
T_{ad}	Adiabatic flame temperature (K)
T_{in}	Inlet (preheat) temperature (K)
T_0	Reference temperature (K)
$\dot{m}_{fu}$	Fuel mass flow rate (kg s^{-1})
$\dot{m}_{air}$	Air mass flow rate (kg s^{-1})
V_{cc}	Combustion chamber volume (m^3)
V_f	Flame volume (m^3)
S_g	Geometric swirl number
ϕ	Global equivalence ratio
ϕ_{max}	Maximum equivalence ratio
$[X]_{eq}$	Equilibrium concentration of X (mol m^{-3})
EINO$_x$	Emission index of NO$_x$ (g kg^{-1})
MW$_{NO}$	Molecular weight of NO (kg mol^{-1})
LHV$_{H_2}$	Lower heating value of hydrogen (kJ kg^{-1})
τ	Mean residence time (s)
T	Random variable of residence time (s)
$P_T(t)$	PDF of residence time T (s^{-1})
σ	Standard deviation of T (s)
$\mathbb{E}[\cdot]$	Expected value operator
K	Overall model constant (lumped parameter)
a	Exponent of fuel mass flow rate
b	Exponent of pressure
c	Exponent of swirl number
d	Exponent of adiabatic flame temperature
α	Residence-time exponent
η	Swirl sensitivity coefficient
γ	Equivalence-ratio sensitivity coefficient

leading to several scaling laws for combustion, such as Constant Velocity (CV) or Constant Residence Time (CRT) approaches, which were developed focusing on geometry and power of the burner [5]. Despite these advances, few scaling laws have been developed for pollutant emissions, particularly for hydrogen combustion. Some laws have been proposed for NO$_x$ in natural-gas systems [5] and, more recently, in hydrogen systems [6–8]. However, some of these scaling laws, such as the one developed by Marragou et al. [6], were derived from experiments in a single combustor configuration, and it remains uncertain whether they remain valid when applied to different burner geometries, swirl intensities, or operating conditions. Consequently, empirical and semi-empirical correlations remain the main predictive tools for NO$_x$ emissions in hydrogen combustion. Among the best known are the correlation by Lefebvre [9], developed from extensive experimental data in aeroengine combustors, and the correlation by Odgers and Kretschmer [10], derived from a rate equation with empirical modifications. Although these correlations are still widely used, they were originally developed for hydrocarbon liquid fuels. Their accuracy, and even validity, for hydrogen combustion is therefore uncertain.

Most modern aeroengines work with swirl-stabilized combustors. Swirl combustion uses guiding vanes or helical blades in a burner to create a rotational motion (swirl) in the fuel and air mixture, leading to improved flame stability and better mixing of reactants. However, the complex flow fields produced by swirlers create strong recirculation zones and multiple flame regimes which make NO$_x$ prediction particularly challenging [11]. Despite their practical relevance, predictive tools specifically developed for swirl-stabilized hydrogen flames are lacking. This gap highlights the need for new scaling laws and analytical correlations tailored to hydrogen-fueled, swirl-stabilized combustors. These models are crucial to the design of more environmentally friendly combustion systems.

In this study, a physics-based, semi-empirical analytical correlation is developed to predict NO$_x$ emissions in partially premixed swirl-stabilized lean H$_2$-air burners, bridging the gap between empirical correlations and purely fundamental scaling laws. To the best of our knowledge, this is the first attempt to develop and validate a physics-based, semi-empirical correlation for NO$_x$ emissions in swirl-stabilized lean hydrogen–air combustion. The model, grounded in physical principles, incorporates semi-empirical corrections and is calibrated with experimental data. A key objective is to provide a tool that is easy to apply in practice, requiring only a minimal set of operating parameters without detailed testing of the burner itself. In this way, the correlation can serve as a plug-and-play first-order prediction method, supporting burner design, experimental planning, and exploration of the parametric space before more detailed investigations are pursued. The derivation of the model, including assumptions and simplifications, is presented in Section 2. Details on the experimental data and pre-processing steps are provided in Section 3, followed by the methodology for fitting the model in Section 4. Results of the model fitting are discussed in Section 5, with further interpretation of the model's accuracy, physical significance, and statistical description in Section 6. The study concludes with a summary of the research outcomes in Section 7.

2. Model derivation

This section presents the step-by-step derivation of an analytical model to predict the emissions index of nitrogen oxides, EINO$_x$ (grams of NO$_x$ produced per kg of fuel burnt), in swirl-stabilized, partially premixed, lean hydrogen–air combustors. The following discussion details this derivation, culminating in a unified expression for EINO$_x$.

The objective of the analytical tool is to establish a relationship between EINO$_x$ and key predictor variables: (1) the fuel mass flow rate $\dot{m}_{fu}$ (which is equivalent to thermal power for a given fuel); (2) the operating pressure of the combustor P_{op}; (3) the adiabatic flame temperature T_{ad} considering the inlet or pre-heating temperature T_{in} and operating pressure P_{op}; (4) the equivalence ratio ϕ; (5) the mean residence time of the combustor τ; and (6) the geometrical swirl number S_g. Other effects, such as, but not limited to, heat losses, detailed injector geometry, or specific fuel–air mixing methods are not considered. This relationship can be expressed in the general functional form of Eq. (1), which will be progressively simplified in the following sections

$$\text{EINO}_x = f\left[\dot{m}_{fu}, P_{op}, T_{ad}(T_{in}, P_{op}, \phi), \tau, S_g\right] \tag{1}$$

The mean residence time τ can be defined as the average duration over which the burnt gases remain in the combustion chamber [12,13]. Considering gas expansion, the average residence time is mathematically expressed in Eq. (2), where V_{cc} is the combustion chamber volume, T_{ad} is the adiabatic flame temperature, and R_{gas} is the specific gas constant. T_{ad} and R_{gas} were computed for a given inlet temperature T_{in}, operating pressure P_{op}, and equivalence ratio ϕ using Cantera (San Diego mechanism with N$_2$ chemistry [14]).

$$\tau = \frac{V_{cc} P_{op}}{(\dot{m}_{air} + \dot{m}_{fu}) T_{ad} R_{gas}} \tag{2}$$

It is acknowledged that in swirl-stabilized combustors, recirculation zones and vortex breakdown cause deviations from an ideal plug-flow assumption. Consequently, the flow fields would be better represented by a residence-time distribution (RTD) rather than a single value. However, previous studies have shown that the mean residence time defined in Eq. (2) provides a reliable approximation of the average of the full distribution at the combustor outlet [15]. The error associated

with neglecting the full distribution is estimated and discussed in Section 5.1.

To simplify the model and reduce its complexity, two main assumptions are made: (1) only NO is considered, as its concentration typically exceeds that of NO_2 and N_2O by orders of magnitude in most practical hydrogen combustion systems [16,17]; and (2) only NO formation via the thermal pathway (Zeldovich mechanism) is accounted for. It is recognized that additional mechanisms may become relevant at elevated pressures [18]. The applicability of the model at elevated pressures is discussed in Section 6.

The derivation considers the total NO production within a homogeneous reactive volume, such that spatially averaged thermochemical quantities represent the flame. The total production rate is obtained by integrating the molar source term over the flame volume and multiplying by the molecular mass of NO, yielding the corresponding mass flow rate. For analytical tractability, a uniform volumetric NO production rate is initially assumed. However, NO formation is inherently spatially non-uniform, it is characterized by peak production in high-temperature flame zones and sustained growth in the post-flame region. This simplification is corrected through the introduction of a factor depending on residence-time, $f(\tau)$, as shown in Eq. (3). The factor $f(\tau)$ provides a first-order representation of post-flame thermal NO growth but does not resolve spatial inhomogeneities in temperature or equivalence ratio. Consequently, the model predicts globally integrated exhaust NO_x rather than spatially distributed NO formation.

For the remainder of this paper, NO will be referred to as NO_x for simplicity.

$$\dot{m}_{NO_x} = \mathrm{MW_{NO}} \int_V [\dot{\mathrm{NO}}_x]\, dv = \mathrm{MW_{NO}}\, [\dot{\mathrm{NO}}_x]\, V_f\, f(\tau) \tag{3}$$

By limiting NO_x formation to only the one produced via the thermal pathway, an Arrhenius-like expression can be used to describe its production rate. This functional form arises from the analytical manipulation of the extended Zeldovich mechanism focused on the rate-limiting step, i.e., $N_2 + O \longrightarrow NO + N$, which leads to Eq. (4), as derived in [19,20].

$$[\dot{\mathrm{NO}}_x] = A T_{\mathrm{ad}}^{-\frac{1}{2}} e^{-\frac{69{,}460}{T_{\mathrm{ad}}}} [\mathrm{O_2}]_{\mathrm{eq}}^{\frac{1}{2}} [\mathrm{N_2}]_{\mathrm{eq}} \tag{4}$$

where A is a constant coefficient, T_{ad} is the adiabatic flame temperature, and $[X]_{\mathrm{eq}}$ is the concentration of molecule X at thermochemical equilibrium conditions. This expression relies on the adiabatic flame temperature, which is assumed to be the flame temperature at given values of T_{in}, P_{op}, and ϕ

Here, equilibrium species concentrations were computed with Cantera (San Diego mechanism with N_2 chemistry [14]), assuming ideal gas and complete chemical equilibrium. A parametric study was conducted to analyze the dependence of $[\mathrm{O_2}]_{\mathrm{eq}}^{1/2}[\mathrm{N_2}]_{\mathrm{eq}}$ on the equivalence ratio ϕ and operating pressure P_{op}. This dependency is shown in Fig. 1(a). The behavior of this quantity can be approximated using an analytical expression. As illustrated in Fig. 1(b), the product $[\mathrm{O_2}]_{\mathrm{eq}}^{1/2}[\mathrm{N_2}]_{\mathrm{eq}}$ can be expressed as: $[\mathrm{O_2}]_{\mathrm{eq}}^{1/2}[\mathrm{N_2}]_{\mathrm{eq}} = K_1 P_{\mathrm{op}}^{b_1} \exp(-\gamma\phi)$, where K_1, b_1, and γ are free parameters fitted to Cantera results using non-linear least squares regression. By incorporating this fitted expression into Eqs. (3) and (4), and combining the constant coefficients A, K_1, and $\mathrm{MW_{NO}}$ into a single constant K, Eq. (5) is obtained. While the pressure exponent b_1 will be discussed later, the equivalence ratio coefficient retains the value determined from regression, with $\gamma = 3.7738$.

$$\dot{m}_{NO_x} = K P_{\mathrm{op}}^{b_1} T_{\mathrm{ad}}^{-\frac{1}{2}} e^{-\left(\frac{69{,}460}{T_{\mathrm{ad}}} + 3.7738\phi\right)} V_f f(\tau) \tag{5}$$

Analytically modeling the flame volume V_f remains challenging, as the relationships between flame morphology and operating parameters are not fully understood for swirling flames, particularly for pure hydrogen combustion. Nevertheless, several phenomenological trends have been reported in the literature for hydrocarbon and hydrogen-enriched flames. A study on flame size, shape, and volume at elevated

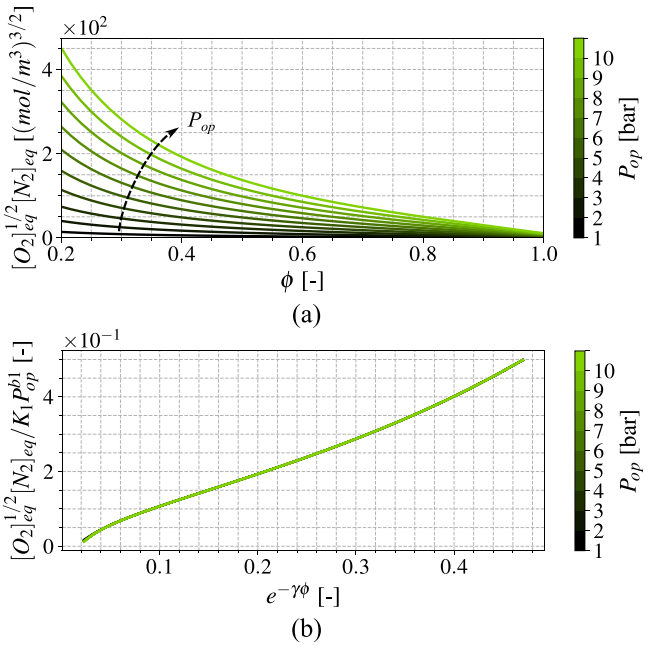

Fig. 1. (a) $[\mathrm{O_2}]_{\mathrm{eq}}^{\frac{1}{2}}[\mathrm{N_2}]_{\mathrm{eq}}$ dependence on ϕ and P_{op}. (b) Fitted analytical expression ($K_1 = 28.671$, $b_1 = 1.5$, $\gamma = 3.7738$).

pressures by Xi et al. [21] for non-premixed methane–air flames showed that both flame length and width increased with fuel mass flow rate $\dot{m}_{\mathrm{fu}}$, while increasing swirl intensity reduced flame length and overall volume. Increasing operating pressure reduced flame length, whereas flame width exhibited a weaker dependence. Flame volume appeared largely independent of injection velocity. Similarly, experiments on a 100 kW swirled burner operating with premixed and non-premixed natural gas by Valera-Medina et al. [22] demonstrated strong sensitivity of flame size to equivalence ratio and swirl intensity, with flame volume decreasing with increasing swirl intensity and equivalence ratio. A weak dependence on Reynolds number was also reported, supporting the limited influence of injection velocity. For hydrogen-enriched methane–air swirling flames, Kim et al. [23] observed that increasing swirl intensity widened the flame while shortening its length, leading to an overall reduction in flame volume.

While these studies primarily concern methane or hydrogen-enriched methane flames, they consistently suggest monotonic dependencies of flame volume on fuel flow rate, pressure, and swirl intensity. Pure hydrogen flames differ in laminar burning velocity, thermo-diffusive behavior, and instability sensitivity; therefore, the numerical scaling exponents cannot be assumed identical. However, the general power-law structure provides a flexible semi-empirical framework capable of capturing these monotonic trends, with hydrogen-specific effects embedded in the fitted exponents.

Based on these phenomenological observations, a scaling for the flame volume of swirled hydrogen flames is proposed as given by Eq. (6). In this equation the flame volume V_f is related through a power law to fuel mass flow rate $\dot{m}_{\mathrm{fu}}$, operating pressure P_{op}, geometric swirl number S_g (which is a simplified approximated description of the swirl intensity) and adiabatic flame temperature T_{ad}

$$V_f \propto \dot{m}_{\mathrm{fu}}^{a_1} P_{\mathrm{op}}^{b_2} S_g^{c} T_{\mathrm{ad}}^{d}, \tag{6}$$

where the exponents a_1, b_2, c, and d are determined through regression using hydrogen combustion data.

Regarding the residence time correction, it has been commonly observed that NO_x emissions follow a linear or power-law trend for residence time [6,10,12,24]. Given its simplicity, the same functional form was adopted, i.e., $f(\tau) = \tau^{\alpha}$.

By combining all the expressions derived so far and dividing the resulting equation by the fuel mass flow rate, an expression for the Emissions Index of NO_x ($EINO_x$) is obtained, as shown in Eq. (7).

$$EINO_x = K\dot{m}_{fu}^{a} P_{op}^{b} S_g^{c} T_{ad}^{d} e^{-\left(\frac{69{,}460}{T_{ad}}+3.7738\phi\right)} \tau^{\alpha} \tag{7}$$

Notably, the pressure exponents from the modeling of O_2 and N_2 equilibrium concentrations (b_1) and from the flame volume (b_2) are merged into a single exponent b. Similarly, the negative half exponent of the adiabatic flame temperature coming from the Arrhenius-like equation is combined with the exponent arising from the flame volume, resulting in a single exponent d. Finally, the mass flow rate exponent a is lumping the flame volume effects (a_1) and the fuel mass flow rate normalization in the definition of $EINO_x$ as $a = a_1 - 1$.

The final step in deriving this model involves incorporating the swirl intensity to account for varying degrees of fuel–air mixing in different burners. As previously mentioned, this model is specifically designed for partially premixed hydrogen–air burners, where the level of mixing can differ between systems. Experimental studies have shown that swirl intensity affects the mixing quality of the injection process, which in turn influences NO_x emissions [7,11,23,25].

Quantifying the degree of mixing analytically remains challenging, and its direct relationship with swirl intensity adds further complexity. In this model, imperfect mixing is accounted for by modifying the global equivalence ratio based on the swirl intensity. The swirl number is a key indicator of the swirl intensity of a flow. However, it is often challenging to measure the effective swirl number precisely. For this reason, it is often approximated by a simplified swirl number based on geometrical features of the swirler [26] commonly known as the geometric swirl number S_g. In this study, the geometrical swirl number S_g is used as a parameter to describe the swirl intensity.

In the model, this effect of imperfect mixing is made explicit: locally richer regions can lead to higher flame temperatures and thus enhanced thermal NO_x production. By incorporating this approach, areas with higher NO_x production are given greater mathematical weight, ensuring that the global effective equivalence ratio more accurately reflects the overall combustion process. This effective equivalence ratio is expressed as the swirl number correction given by Eq. (8), where ϕ_{eff} represents the global effective equivalence ratio, which is the adjusted equivalence ratio that accounts for mixing by assigning greater weight to richer zones. In this formulation, ϕ is the nominal global equivalence ratio, while ϕ_{max} and η are free parameters. Specifically, ϕ_{max} defines the upper bound of ϕ_{eff}, and η controls the sensitivity of ϕ_{eff} to variations in swirl intensity, as characterized by the geometric swirl number.

$$\phi_{eff} = (\phi_{max} - \phi)e^{-\eta S_g} + \phi \tag{8}$$

The proposed formulation is phenomenological and is constructed to satisfy two physically consistent asymptotic limits. In the limit of vanishing swirl intensity ($S_g \to 0$), mixing is minimal and the effective equivalence ratio approaches an upper bound ϕ_{max}, representing maximum mixture stratification. In the opposite limit of very high swirl intensity ($S_g \to \infty$), enhanced mixing leads ϕ_{eff} to approach the nominal global equivalence ratio ϕ, corresponding to perfect premixing. The exponential form ensures a smooth and monotonic transition between these limits and captures the diminishing sensitivity of mixing enhancement at higher swirl numbers.

The parameter ϕ_{max} should therefore be interpreted as a phenomenological upper asymptote within the model structure rather than a physically realized flame condition. The majority of the swirled flames included in the dataset have $S_g \geq 0.6$. Therefore, the $S_g \to 0$ limit lies outside the calibration range, and ϕ_{max} represents an extrapolated bound rather than a directly observable operating state.

When using the effective global equivalence ratio, all quantities that originally depended on the equivalence ratio must be adjusted accordingly, namely, the adiabatic flame temperature T_{ad} and the residence time τ. First, the adiabatic flame temperature is recalculated based on the effective global equivalence ratio: $T_{ad,eff}(T_{in}, P_{op}, \phi_{eff})$. Then, the residence time is modified as described in Eq. (9).

$$\tau_{eff} = \tau \frac{T_{ad}(\phi) R_{gas}(\phi)}{T_{ad,eff} R_{gas}(\phi_{eff})} \tag{9}$$

With these adjustments, the final form of the model is given by Eq. (10) together with Eq. (8). The final form requires 6 predictor variables and it has 8 free parameters ($K, a, b, c, d, \alpha, \eta, \phi_{max}$) that could be adjusted to experimental data.

$$EINO_x = K\dot{m}_{fu}^{a} P_{op}^{b} S_g^{c} T_{ad,\ eff}^{d} e^{-\left(\frac{69{,}460}{T_{ad,eff}}+3.7738\phi_{eff}\right)} \tau_{eff}^{\alpha} \tag{10}$$

The constant K serves as a dimensional balancing parameter absorbing unresolved physical effects. Its compound dimensions ensure correct output units but are not shown explicitly in figures. Quantities involving K are therefore reported in balanced units (b.u.).

3. Data mining

Publicly available NO_x emission data for swirl-stabilized partially-premixed hydrogen–air burners are scarce. Despite this, a dataset was compiled from seven burners to determine the model's free parameters for H_2-air swirl-stabilized combustion. The analyzed studies include: (1) a 12 kW combustor using a dual-swirl dual-fuel (DSDF) injector, considering only pure hydrogen data [27]; (2) HYLON burner tests with a DSDF injector at constant equivalence ratios and thermal power [13]; (3) HYLON burner experiments with varying pre-heating temperatures [12]; (4) high-pressure HYLON burner tests with the DSDF injector [6]; (5) a 10 kW combustor study with a dual swirler, testing different injector geometries and swirl intensities [7]; (6) a 100 kW combustor with a radial swirler and axial air injection [25]; and (7) a 12 kW combustor with an axial swirler and axial air injection [11]. The final dataset comprises 262 points and a summary of the operating conditions ranges for each dataset is given in Table 1.

NO_x emission data from the literature required pre-processing for consistency. Values were reported either as Emission Index ($EINO_x$) or in parts per million (ppm) with varying oxygen reference percentages, in which case they were converted to $EINO_x$ for uniformity.

Most predictor variables were directly extracted from the literature, except for the geometric swirl number, which varies based on different formulations [26]. For double-swirl injectors, an effective swirl number was calculated by weighting each swirler's geometric swirl number according to its axial momentum flux. Some burners used axial air injection, which affects swirl intensity, but since no analytical models exist to account for this effect, the geometric swirl number was used unmodified, potentially introducing prediction errors. The complete dataset and preprocessing scripts are available as described in Section "Data availability".

4. Model fitting methodology

An optimization procedure was performed to determine the free parameters that best fit the model to experimental data. The cost function minimized the squared errors between the predicted and measured $EINO_x$ values, normalized by the number of measurements n, as shown in Eq. (11), where y_i represents the measured $EINO_x$ value, and $EINO_{x,i}(\bar{x})$ is the predicted value for the corresponding vector of free parameters $\bar{x} = [K, a, b, c, d, \alpha, \eta, \phi_{max}]$.

$$\min_{\bar{x}} f(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n}\left(EINO_{x,i}(\bar{x}) - y_i\right)^2 \tag{11}$$

The optimization used a differential evolution algorithm with dithering for fast convergence. The mutation constant ranged from 0.5 to 1, and the recombination constant was 0.7. A population size dependence study led to a population size of 100. The DE/best/1/bin strategy was

Table 1
Ranges of operating conditions for the gathered data.

Study	Power [kW]	Equivalence Ratio [-]	Pre-Heat Temperature [K]	Pressure [bar]	Swirl Number [-]	# Datapoints
Marragou, et. al. 2022 [27]	3.9–12.9	0.45–0.62	293	1	0.67	12
Marragou 2023 [13]	3.6–12.7	0.2–0.8	293	1	0.67	30
Magnes, et. al. 2023 [12]	2.0–21.0	0.3–0.4	293–673	1	0.67	52
Marragou, et. al. 2025 [6]	5.7–51.2	0.2–0.6	293	2–8	0.67	97
Leroy, et. al. 2024 [7]	7.5–12.4	0.35	293	1	0.02–0.9	29
Reichel, et. al. 2018 [25]	87.4–122.4	0.5–0.7	450–620	1	0.7–0.9	29
Link, et. al. 2025 [11]	7.0–13.0	0.39–0.73	298	1	0.7–1.10	13

chosen for its strong performance [28]. Parameter constraints, given in Eq. (12), ensured flexibility while preventing unphysical solutions, with some bounds derived from previous NO_x correlations [29–31]

$$\begin{aligned} -3 \le a \le 3 \qquad & 10^{10} \le K \le 10^{30} \\ -3 \le b \le 3 \qquad & 0 \le \alpha \le 5 \\ -3 \le c \le 3 \qquad & 0 \le \eta \le 10 \\ -5 \le d \le 5 \qquad & 0.8 \le \phi_{\max} \le 2.0 \end{aligned} \tag{12}$$

To stabilize optimization, terms in Eq. (10) influenced by the free parameters a, b, and d were normalized to balance the model's sensitivity as shown in Eq. (13). The fuel mass flow rate was scaled by the hydrogen low heating value, $\mathrm{LHV}_{H_2} = 119.96 \times 10^3$ kJ/kg, making thermal power in kW the driving parameter. Operating pressure and the (effective) adiabatic flame temperature were normalized to ambient conditions ($P_0 = 1$ bar, $T_0 = 298$ K).

$$\mathrm{EINO}_x = K(\dot{m}_{\mathrm{fu}}\mathrm{LHV}_{H_2})^a \left(\frac{P_{\mathrm{op}}}{P_0}\right)^b S_g^c \left(\frac{T_{\mathrm{ad,\ eff}}}{T_0}\right)^d e^{-\left(\frac{69{,}460}{T_{\mathrm{ad,eff}}} + 3.7738\phi_{\mathrm{eff}}\right)} \tau_{\mathrm{eff}}^{\alpha} \tag{13}$$

To validate the model, the dataset was split into 75% training and 25% validation subsets. To ensure that the training and validation subsets captured the full variability of the operating space, k-means clustering was applied to the predictor variables and burner identity. This approach groups data into clusters based on similarity. The number of clusters is given by the size of the training sample, leaving the remainder for validation. Compared to random splitting, this approach reduces over-representation of densely sampled regions (e.g., atmospheric pressure or same burner) and ensures that less frequently reported but relevant operating regimes are represented in both subsets. As a result, the fitted parameters are exposed to a broader and more balanced range of flame conditions, improving robustness and limiting overfitting to localized portions of the dataset. Nevertheless, the clustering strategy cannot fully compensate for the uneven availability of literature data. Burners with a larger number of reported operating points inevitably contribute more samples, which may influence the fitted correlations. This is acknowledged as an inherent limitation of the present study. The data and the model are publicly available (see Section "Data availability"). Therefore, if new data is released in the literature the model can be re-calibrated to accommodate for additional burner configurations.

5. Model performance and uncertainty

This section evaluates the model's performance in predicting EINO_x values and quantifies the associated uncertainties. To determine the optimal free parameters and their uncertainties, the optimization procedure was run 1000 times, storing the free parameter vector from each run for statistical analysis. A typical optimization run yields an average predictive error of ±0.57(0.69) g/kg or ±53.1(46.9)% in EINO_x prediction, with a determination coefficient of $R^2 = 0.44(0.74)$ for validation (training). Fig. 3.a shows the parity plot of measured versus predicted EINO_x values for the training dataset, while Fig. 3.b shows the corresponding plot for the validation dataset. To assess the stability of the fitting procedure, the distributions of the resulting R^2 scores and Mean Absolute Error (MAE) values over all 1000 optimization runs are shown in Fig. 2. The histograms exhibit a very narrow spread, indicating that the optimization algorithm consistently converges to nearly identical objective-function. The limited variability in both R^2 and MAE confirms that the fitting process is robust and that the predictive performance is not sensitive to initialization or local minima.

Table 2
Statistics of the free parameter values.

	16th Percentile	Median (50th)	84th Percentile
K	$10^{18.819957}$	$10^{19.174242}$	$10^{19.468171}$
a	0.340918	0.341781	0.342289
b	0.379393	0.380467	0.382460
c	−0.618658	−0.616796	−0.612322
d	−5.000000	−4.673504	−4.275361
α	0.736562	0.741203	0.746653
η	0.447910	0.454375	0.458263
$\phi_{\max}$	1.157742	1.159688	1.160024

Table 2 presents the obtained free parameter values and uncertainties, expressed as percentiles (16-50-84) due to their asymmetric distribution.

To estimate uncertainties in EINO_x predictions, a Monte Carlo simulation with 10,000 samples was performed on the entire dataset. The mean and standard deviation were computed to obtain error bars for each data point. Since some free parameters exhibited high correlation and no clear parametric distribution, Gaussian Kernel Density Estimation (KDE) was used for Monte Carlo sampling. The resulting prediction uncertainties are shown in Fig. 3.c.

5.1. Mean residence time error estimation

As discussed earlier, the model developed in this study prioritizes simplicity and practical applicability. However, it remains important to quantify the error introduced by approximating the Residence Time Distribution (RTD) with a single mean value.

The mean residence time defined in Eq. (2) has been shown to provide a reliable approximation of the average of the RTD at the combustor outlet [15]. Let T be a random variable describing the RTD with probability density function $P_T(t)$. Then the mean residence time is the expected value, $\tau = \mathbb{E}[T] = \int_{-\infty}^{\infty} t P_T(t)\, dt$, while the variance is $\sigma^2 = \mathrm{Var}(T) = \int_{-\infty}^{\infty} (t-\tau)^2 P_T(t)\, dt$.

In the present model, the RTD is considered only through the correction factor $f(\tau) = \tau^{\alpha}$. Using the mean residence time thus implies $f(\tau) = \mathbb{E}[T]^{\alpha}$, whereas the more accurate formulation would be given by Eq. (14), which accounts for the full RTD.

$$\mathbb{E}[T^{\alpha}] = \int_{-\infty}^{\infty} t^{\alpha} P_T(t)\, dt \tag{14}$$

To estimate the error introduced by replacing Eq. (14) with $f(\tau)$, a second-order Taylor expansion of $f(T)$ around τ is considered, as shown in Eq. (15).

$$f(T) \approx f(\tau) + f'(\tau)(T-\tau) + \frac{1}{2} f''(\tau)(T-\tau)^2 \tag{15}$$

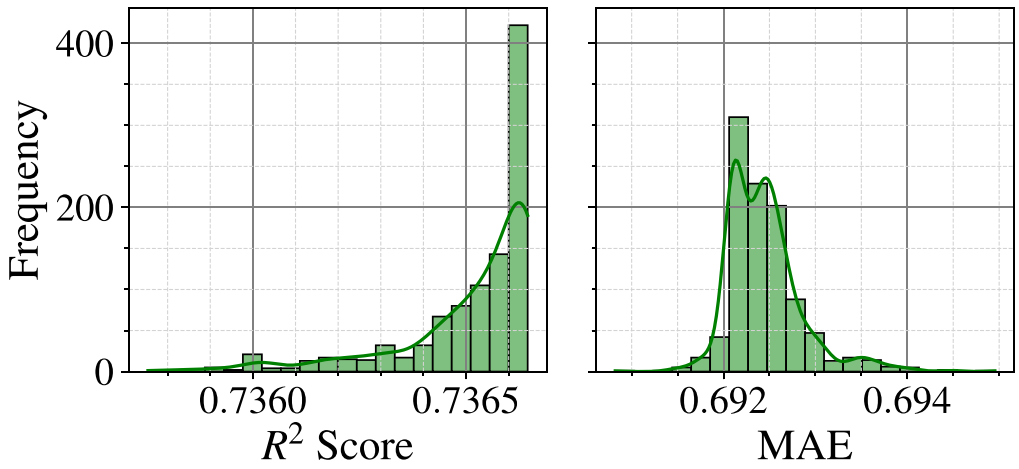

Fig. 2. Histograms of R^2 score and Mean Absolute Error (MAE).

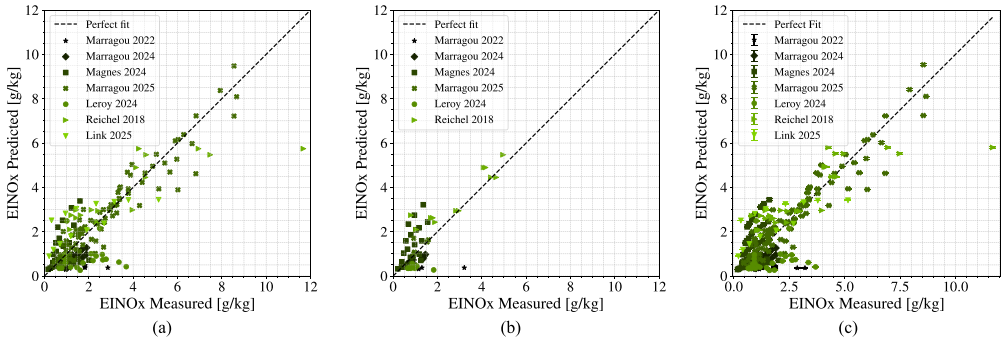

Fig. 3. (a) Model performance on the training dataset (Error: ±0.69 g/kg, R^2 = 0.74). (b) Model performance on the validation dataset (Error: ±0.57 g/kg, R^2 = 0.44). (c) Predicted vs. measured $EINO_x$ considering uncertainty from the free parameters.

Taking expectations yields Eq. (16).

$$\mathbb{E}[f(T)] \approx \tau^{\alpha} + \frac{1}{2} f''(\tau)\sigma^2 \tag{16}$$

Since $f''(\tau) = \alpha(\alpha-1)\tau^{\alpha-2}$, Eq. (16) becomes Eq. (17).

$$\mathbb{E}[f(T)] \approx \tau^{\alpha} + \frac{1}{2}\alpha(\alpha-1)\tau^{\alpha-2}\sigma^2 \tag{17}$$

The relative error (R.E.) introduced by the mean-value approximation is then expressed by Eq. (18).

$$\text{R.E.} = \frac{\mathbb{E}[T^{\alpha}] - \mathbb{E}[T]^{\alpha}}{\mathbb{E}[T]^{\alpha}} \approx \frac{1}{2}\alpha(\alpha-1)\left(\frac{\sigma}{\tau}\right)^2 \tag{18}$$

For the present model with $\alpha = 0.74$, the factor $\alpha(\alpha-1)$ is negative, implying R.E. < 0. Thus, the mean-value approximation systematically overestimates the influence of residence time. The magnitude of this conservatism scales with $(\sigma/\tau)^2$, and therefore depends on the width of the RTD.

Reported numerical and experimental RTD measurements in swirl combustion systems show $(\sigma/\tau)^2$ values typically in the range of 0.4–0.85 [32–34]. Substituting this range into Eq. (18) yields predicted relative errors between approximately 4% and 8%.

These results indicate that the mean-value approximation introduces a moderate but bounded conservatism. Given the overall modeling uncertainties and the global nature of the present formulation, this level of error remains acceptable for the intended predictive scope.

6. Discussion

This section evaluates the model's quality and performance. Section 6.1 discusses accuracy and compares it to similar models. The physical significance of the free parameters and their uncertainties is examined in Section 6.2, while Section 6.3 addresses statistical properties and their impact on model performance.

6.1. Model accuracy

As shown in Fig. 3.c, despite uncertainties in the optimized free parameters, the propagation of this uncertainty into the predictions is relatively small, as indicated by the narrow error bars. This suggests that the model achieves high precision, even though its overall accuracy remains moderate. Predicting NO_x is inherently challenging, with reported errors exceeding 100% in some models and datasets [9,10,29–31]. When limited information about combustor geometry or operating conditions is available, predictions often provide only order-of-magnitude estimates [35]. For instance, an RMS error of 173% has been reported in a gas turbine correlation [10]. Even widely used correlations demonstrate significant errors, particularly when applied outside their original development range [9]. More accurate approaches exist, such as the scaling law of Marragou et al. [6] (R^2 = 0.91) and the correlation of Wang et al. [36] (R^2 = 0.96), but their performance is strongly tied to the specific combustor configurations and operating conditions on which they were developed. In this context, while the present model shows only moderate accuracy, its performance is comparable to or better than several established models.

It is worth noting the pronounced prediction scatter at low NO_x levels in Fig. 3. Such behavior is common across correlations due to the inherent difficulty of estimating emissions in this regime. Several factors may contribute: (1) low emissions approach the detection limits of measurement instruments, increasing relative uncertainty; (2) in low-emission regimes, multiple formation pathways may contribute with similar importance, in contrast to high-load conditions where thermal

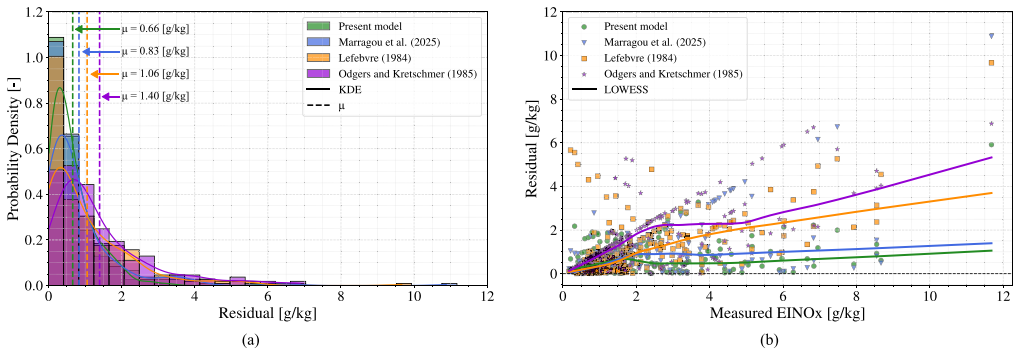

Fig. 4. (a) Probability density of residuals obtained via Kernel Density Estimation (KDE), with dashed vertical lines marking mean residuals for different models in the literature. (b) Residuals plotted against experimentally measured $EINO_x$ values, with LOWESS curves showing residual trends. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

NO dominates. This complicates prediction based on a "thermal-only" assumption.; and (3) prediction scatter at low emissions appears more pronounced on linear scales, while logarithmic scaling, which is commonly used in other studies, tends to compress high values and reduce its visual impact.

To assess accuracy quantitatively, the model was benchmarked against three references: two widely used gas turbine correlations (Lefebvre [9] and Odgers and Kretschmer [10]) and the recent scaling law of Marragou et al. [6]. For a fair comparison, each model was re-calibrated to the present dataset by adjusting the multiplicative constant in its formulation. Since the Lefebvre and Odgers correlations were originally expressed in terms of $EINO_x$, no further modification was needed; however, Marragou's model was developed in terms of concentration (ppm) and was therefore converted to $EINO_x$ for consistency.

Residuals were computed for each model using the same dataset as in the present study. The probability density functions of these residuals, obtained using Kernel Density Estimation (KDE), are presented in Fig. 4.a. The Lefebvre and Odgers correlations yield higher mean residuals compared to both Marragou's scaling law and the present model. Moreover, the distributions of Marragou's and the current model's residuals are more sharply centered around lower values, indicating a greater ability to predict NO_x with reduced error. The mean residual for the model of this study is lower than that of Marragou et al. potentially reflecting the fact that the present model explicitly accounts for mixing effects, which were neglected in their study.

A complementary perspective is given in Fig. 4.b, where the residuals of each datapoint are plotted against the experimentally measured values of $EINO_x$. LOWESS (LOcally WEighted Scatterplot Smoothing) curves highlight the trends of residuals across different emission levels. The results show that the Lefebvre and Odgers correlations exhibit residuals that grow in magnitude with increasing NO_x, reflecting a systematic bias at higher emission levels. By contrast, both Marragou's scaling law and the present model display residuals that converge asymptotically toward lower values, suggesting more uniform predictive capability across a broad range of operating conditions.

Taken together, these results demonstrate that the proposed model is superior to traditional correlations and, in some cases, performs on par with state-of-the-art models, such as that developed by Marragou et al. [6]. While both models achieve similar accuracy, the present model is applicable to a wider variety of burners and is not limited to a single configuration. Its robustness across a range of conditions underscores its potential for practical use, where both generality and reliable accuracy are required.

An additional note must be made on the applicability of this model. Although the present formulation assumes thermal NO formation as the governing mechanism, detailed kinetic studies of lean hydrogen and hydrocarbon combustion [37–39] indicate that the N_2O-intermediate pathway contributes non-negligibly even at atmospheric pressure and increases markedly with pressure up to approximately 6–10 bar. Beyond this range, the N_2O pathway often becomes dominant, although its relative contribution tends to plateau with further pressure increase.

Since the semi-empirical model is calibrated using literature data spanning atmospheric conditions up to approximately 8 bar, it inherently captures much of this transition regime in which the N_2O contribution rises rapidly with pressure. Consequently, within the investigated pressure range, the observed pressure sensitivity of total NO is effectively embedded in the fitted correlations, even though individual pathways are not explicitly resolved.

At pressures substantially exceeding the calibration range, where the N_2O pathway may dominate, the model may exhibit systematic deviations. However, because the parameters were derived from total NO data rather than pathway-resolved kinetics, the isolated contribution of the N_2O route cannot be separated and a quantitative error specifically attributable to this mechanism cannot be directly estimated. The magnitude of any deviation would depend on how strongly the pressure scaling of NO formation departs from that represented in the calibration dataset and therefore cannot be rigorously quantified without additional high-pressure validation data which is currently scarce in the literature.

6.2. Physical significance of the free parameters

Providing a physical interpretation of the free parameters derived from the optimization process is essential to understand the fundamental mechanisms that contribute to NO_x production and their relative importance.

The fuel mass flow rate exponent, $a = 0.38$, suggests that $EINO_x$ has low sensitivity to changes in thermal power. This exponent represents both the flame volume and the normalization by fuel mass flow rate to align with the $EINO_x$ definition. From the flame volume perspective, $a_1 = a + 1 = 1.38$ (see Eq. (6)), indicates that flame volume increases with fuel mass flow rate. This aligns with [21], who observed a nearly linear increase in flame volume with fuel mass flow rate.

The operating pressure exponent, $b = 0.38$, shows that NO_x production is slightly sensitive to pressure changes. This exponent captures two effects: one related to the equilibrium of O_2 and N_2 molecules (see Fig. 1), affecting the chemical reaction rate in Eq. (4), and another related to flame volume (see Eq. (6)), so that $b = b_1 + b_2$. The fitted expression of the O_2 and N_2 equilibrium effect (see Fig. 1) yielded $b_1 \approx 1.50$, implying that the pure effect of flame volume is $b_2 \approx -1.12$. This finding matches phenomenological descriptions from [21], who noted that flame volume decreases with pressure. Thus, NO_x is positively affected by pressure through increased O_2 and N_2 concentrations ($b_1 > 0$) but negatively affected by the reduced flame volume ($b_2 < 0$).

The residence time exponent, $\alpha = 0.74$, is consistent with previous studies that show a linear relationship between NO_x and residence

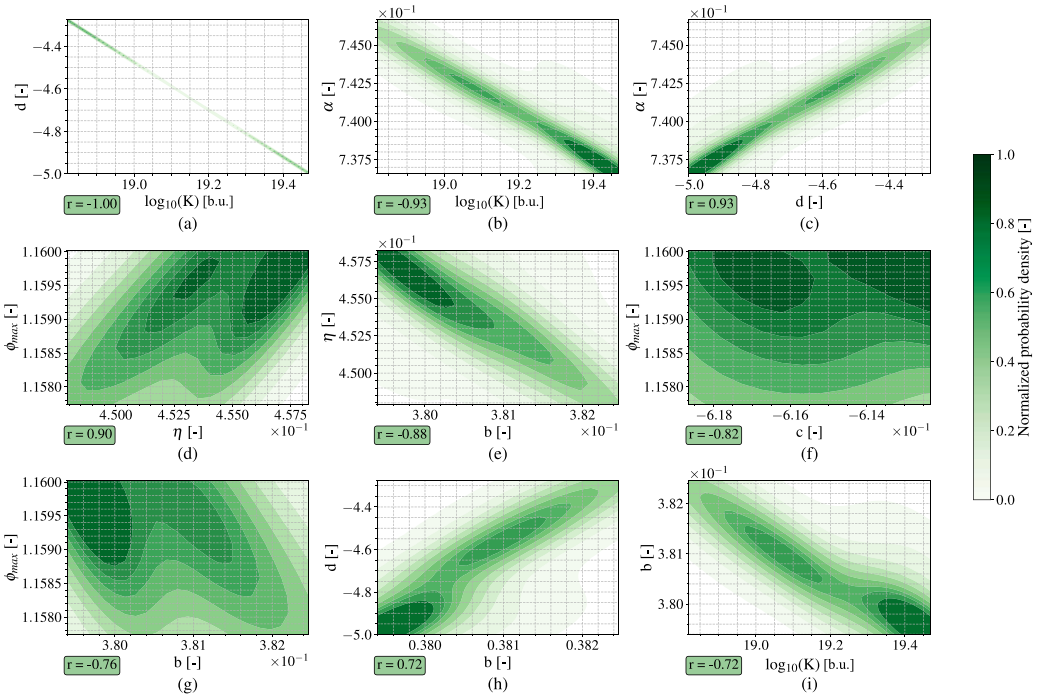

Fig. 5. Joint probability density distributions of highly correlated parameter pairs ($r > 0.7$), obtained using Gaussian Kernel Density Estimation (KDE) normalized with their respective maximum value. Each subfigure (a–i) represents a pair of free parameters, with their possible values on the x- and y-axes, constrained within the 16th and 84th percentiles. The Pearson correlation coefficient (r) for each pair is displayed in a box below the corresponding subfigure.

time [40]. The value is also similar to the reported exponent for residence time in [6,12], which has a value of 1/3. Differences in the definition of residence time and inclusion of mixing effects may explain the variation in value, but the general agreement in sign and order of magnitude supports the results and well known behavior of NO_x with respect to residence time.

The swirl number exponent, $c = -0.61$, indicates that NO_x emissions decrease with swirl intensity. This also suggests that flame volume is inversely proportional to swirl intensity, consistent with previous studies [21–23]. While NO_x is known to be temperature-dependent [41], this study includes an additional effect through flame volume (see Eq. (6)), reflected in the exponent $d = -4.67$. However, this parameter has the highest uncertainty among the free parameters, which may limit its physical interpretation, as discussed in Section 6.3.

Regarding the swirl intensity correction, the maximum effective equivalence ratio, $\phi_{max} = 1.15$, and the swirl sensitivity parameter, $\eta = 0.45$, were obtained. The parameter ϕ_{max} should be interpreted as a phenomenological upper asymptote representing the maximum degree of mixture stratification within the model structure. It reflects the upper bound of local enrichment effects that increase thermal NO_x formation.

From Eq. (8), increasing swirl intensity progressively reduces the effective enrichment, driving ϕ_{eff} toward the nominal global equivalence ratio ϕ. This behavior is consistent with the physical expectation that stronger swirl enhances mixing and reduces large-scale stratification [42]. The moderate value of η indicates that this reduction occurs gradually, implying that even at elevated swirl numbers some degree of mixture inhomogeneity persists. This observation aligns with the partial-premixing hydrogen injection strategy used in the burners considered.

The constant multiplier K at the front of the model is challenging to interpret, as it absorbs all neglected effects, likely making it a significant source of error in the predictions.

6.3. Statistical properties of the free parameters

Beyond their physical interpretation, it is important to assess the statistical properties of the optimized free parameters, since their distribution directly affects uncertainty propagation, parameter identifiability, and the robustness of the model calibration. In particular, testing for normality provides a first indication of whether standard confidence intervals and error estimates are appropriate, or whether the parameter space is skewed or multimodal, which could suggest overparameterization or compensation effects between variables. Although this analysis is not central to the main objectives of the study, it offers complementary insight into the stability and reliability of the parameter set obtained.

A statistical analysis was conducted on the free parameters obtained from 1000 optimization runs to evaluate the robustness and interdependence of the fitted coefficients. The conditions of the optimization runs are given in Section 4. To assess whether conventional uncertainty metrics (e.g. mean and standard deviation) adequately describe the parameter space, a test for normality was conducted. The Shapiro–Wilk test was applied for univariate distributions, while multivariate normality was tested by examining whether the Mahalanobis distances followed a χ^2 distribution [43]. Both tests indicated deviations from normality, implying that the parameter space is structured and potentially shaped by strong correlations rather than random variation. A strong correlation was observed between certain parameter pairs, as determined from their Pearson correlation matrix (not shown here). To better capture these structured relationships, joint probability distributions were generated using Gaussian Kernel Density Estimation (KDE) for parameter pairs with Pearson correlation coefficient values exceeding $r > 0.7$. The results are presented in Fig. 5, where each subfigure (a–i) displays a joint KDE plot for a specific parameter pair. The x- and y-axes represent the range of each parameter, constrained within their 16th and 84th percentiles. Below each plot, the Pearson correlation coefficient quantifies the strength of the correlation.

Several key observations can be drawn from Fig. 5. A strong negative correlation is evident between K and d (Fig. 5.a), K and α (Fig.

5.b) and, to a lesser extent, K and b (Fig. 5.i). Notably, the correlation between K and d, the global constant coefficient and the (effective) adiabatic flame temperature exponent, respectively, stands out. Their density distribution is both narrow and diagonal, indicating an almost perfect inverse relationship.

This behavior can be understood by considering the logarithmic form of the model. The temperature contribution can be written as in Eq. (19)

$$\log_{10} \mathrm{EINO}_x = K + d \log_{10}\left(\frac{T_{\mathrm{ad,eff}}}{T_0}\right) + \ldots \tag{19}$$

where K acts as an intercept and d as a slope parameter. For models of the type $y = K + dx$, regression theory predicts strong negative intercept–slope correlation when the predictor x has a large nonzero mean relative to its variance. In the present dataset, the normalized temperature T/T_0 spans a narrow interval, resulting in $\log_{10}(T/T_0)$ having a substantial positive mean compared to its standard deviation. Under these conditions, variations in d produce nearly uniform shifts in the model response, which are compensated by opposite variations in K, leading to the observed near-perfect inverse correlation.

To evaluate whether this indicates parameter redundancy, the model was re-assessed with $d = 0$. This simplification led to a substantial deterioration in predictive accuracy, demonstrating that the temperature exponent is structurally required. Therefore, the strong K–d correlation reflects geometric intercept–slope coupling rather than over-parameterization.

Furthermore, Fig. 5.c and 5.h show correlations between d and the residence time exponent α, and between d and the pressure exponent b. However, these correlations occur within a relatively narrow uncertainty band and do not lead to significant degradation in predictive capability. They are thus interpreted as practical parameter interactions rather than evidence of structural redundancy.

Another notable observation is the relationship between $\phi_{\max}$ and other parameters. This parameter exhibits high correlation with η (Fig. 5.d), the geometric swirl number exponent c (Fig. 5.f), and the pressure exponent b (Fig. 5.g), indicating its strong coupling to these parameters. Unlike the previously discussed case, these correlations are weaker, and the associated parameter uncertainties are relatively low. The interdependence of $\phi_{\max}$ with multiple parameters reflects the model's flexibility in capturing several physical effects simultaneously, due to the chosen analytical form for swirl intensity correction and the artificial nature of $\phi_{\max}$. While this suggests opportunities for exploring alternative formulations to potentially simplify the representation, it also underscores the model's capability to accommodate complex interactions in swirl combustion.

7. Conclusion

This study presents a semi-empirical analytical model for predicting NO_x emissions in partially-premixed, swirl-stabilized lean hydrogen combustors. The main findings and contributions can be summarized as follows:

- Unlike existing correlations largely developed for hydrocarbon or non-swirling flames, the present model is calibrated exclusively on hydrogen combustion data. The resulting scaling exponents therefore reflect hydrogen-specific flame characteristics.
- A phenomenological swirl correction is introduced to account for imperfect premixing through an effective equivalence ratio formulation. While semi-empirical, this approach provides a physically interpretable mechanism to represent how local near-stoichiometric zones can dominate thermal NO_x formation. Such a correction is absent in traditional non-swirled correlations.
- The model captures the governing mechanisms of NO_x production:
  - Fuel mass flow rate exhibits a weak influence.
  - Pressure affects emissions through competing mechanisms (enhanced reactant concentrations versus reduced flame volume).
  - Increased swirl intensity suppresses NO_x, consistent with enhanced mixing and reduced effective flame volume.
  - Residence time strongly influences NO_x production.
- Although the mean relative error is approximately 50%, the absolute prediction errors remain within a practically relevant range across multiple hydrogen burner designs, demonstrating robustness despite dataset heterogeneity.
- In contrast to CFD-based simulations, chemical reactor network models, or data-driven approaches, which require detailed geometric, flow-field, or measured thermochemical inputs, the present formulation relies solely on global operating parameters. This enables rapid estimation of NO_x emissions prior to combustor fabrication or high-fidelity numerical analysis.
- Future work should aim to refine the swirl-mixing representation and validate its structure through targeted experimental investigations. Expanding high-quality hydrogen emission datasets will further strengthen predictive accuracy.

CRediT authorship contribution statement

Alam Garcidueñas-Correa: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Francesca De Domenico:** Writing – review & editing, Supervision, Conceptualization. **Arvind Gangoli Rao:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alam Garcidueñas-Correa reports financial support was provided by Horizon Europe. Alam Garcidueñas-Correa reports financial support was provided by UK Research and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has received funding from the Horizon Europe Research and Innovation Programme under GA n° 101096275. This work is co-funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee n° 10068673.

The author also acknowledges the support of the CYPHER COST Action (CA22151) to present this study at the 12th European Combustion Meeting (ECM).

Data availability

The full NO_x emissions dataset as well as the Python codes to calibrate the model can be found under the following DOI: 10.4121/8df3f562-0ded-4308-89a6-9b3b8a5c1a68.

References

[1] Habib MA, Abdulrahman GA, Alquaity AB, Qasem NA. Hydrogen combustion, production, and applications: A review. Alex Eng J 2024;100:182–207. http://dx.doi.org/10.1016/j.aej.2024.05.030.

[2] Le TT, Sharma P, Bora BJ, Tran VD, Truong TH, Le HC, Nguyen PQP. Fueling the future: A comprehensive review of hydrogen energy systems and their challenges. Int J Hydrog Energy 2024;54:791–816. http://dx.doi.org/10.1016/j.ijhydene.2023.08.044.

[3] Lewis AC. Optimising air quality co-benefits in a hydrogen economy: a case for hydrogen-specific standards for NOx emissions. Environ Sci: Atmospheres 2021;1(5):201–7. http://dx.doi.org/10.1039/d1ea00037c.

[4] SPALDING D. The art of partial modeling. In: Ninth symposium (international) on combustion. Elsevier; 1963, p. 833–43. http://dx.doi.org/10.1016/b978-1-4832-2759-7.50092-3.

[5] Drubetskoi E, Eckart S, Krause H. Short overview on combustion systems scale-up with emphasis on NOx emissions of gas-fired furnaces. Energy Sci Eng 2021;10(2):621–9. http://dx.doi.org/10.1002/ese3.1028.

[6] Marragou S, Mengu D, Es-sebbar E-t, Magnes H, Schuller T, Guiberti TF. Flame stabilization and pollutant emissions from a H2/air dual swirl coaxial injector at elevated pressure. Int J Hydrog Energy 2025;100:163–72. http://dx.doi.org/10.1016/j.ijhydene.2024.12.133.

[7] Leroy M, Puggelli S, Mirat C, Renaud A, Leparoux J, Buisson Q, Mercier R, Vicquelin R. Investigation of NOx scaling laws in swirled partially premixed hydrogen flames on a coaxial injector. Int J Hydrog Energy 2024;91:256–66. http://dx.doi.org/10.1016/j.ijhydene.2024.10.016.

[8] Meraner C, Li T, Ditaranto M, Løvås T. Effects of scaling laws on the combustion and NO characteristics of hydrogen burners. Combust Flame 2020;214:407–18. http://dx.doi.org/10.1016/j.combustflame.2020.01.010.

[9] Lefebvre AH. Fuel effects on gas turbine combustion-liner temperature, pattern factor, and pollutant emissions. J Aircr 1984;21(11):887–98. http://dx.doi.org/10.2514/3.45059.

[10] Odgers J, Kretschmer D. The prediction of thermal NOx in gas turbines. In: Volume 2: coal, biomass and alternative fuels; combustion and fuels. IGT1985, American Society of Mechanical Engineers; 1985, http://dx.doi.org/10.1115/85-igt-126.

[11] Link S, Dave K, de Domenico F, Rao AG, Eitelberg G. Experimental analysis of dual-fuel (CH4/H2) capability in a partially-premixed swirl stabilized combustor. Int J Hydrog Energy 2025;101:427–37. http://dx.doi.org/10.1016/j.ijhydene.2024.12.286.

[12] Magnes H, Marragou S, Aniello A, Selle L, Poinsot T, Schuller T. Impact of preheating on flame stabilization and NOx emissions from a dual swirl hydrogen injector. J Eng Gas Turbines Power 2023;146(5). http://dx.doi.org/10.1115/1.4063719.

[13] Marragou S. flow structure, mixing, flame stabilization and pollutant emissions from a coaxial dual swirl CH4/H2/air injector [Ph.D. thesis], (2023INPT0031). Institut National Polytechnique de Toulouse - INPT; 2023.

[14] University of California at San Diego, Combustion Research Group. Chemical-Kinetic Mechanisms for Combustion Applications, nSan Diego Mechanism web page, Mechanical and Aerospace Engineering (Combustion Research), https://web.eng.ucsd.edu/mae/groups/combustion/mechanism.html.

[15] Goeckeler K, Terhaar S, Lacarelle A, Paschereit C. Residence time distribution in a swirl-stabilized combustor at cold conditions. In: 41st AIAA fluid dynamics conference and exhibit. American Institute of Aeronautics and Astronautics; 2011, http://dx.doi.org/10.2514/6.2011-3585.

[16] Shudo T, Omori K, Hiyama O. NO reduction and NO2 emission characteristics in rich-lean combustion of hydrogen. Int J Hydrog Energy 2008;33(17):4689–93. http://dx.doi.org/10.1016/j.ijhydene.2008.05.034.

[17] ILBAS M, YILMAZ I, KAPLAN Y. Investigations of hydrogen and hydrogen–hydrocarbon composite fuel combustion and emission characteristics in a model combustor. Int J Hydrog Energy 2005;30(10):1139–47. http://dx.doi.org/10.1016/j.ijhydene.2004.10.016.

[18] Capurso T, Laera D, Riber E, Cuenot B. NOx pathways in lean partially premixed swirling H2-air turbulent flame. Combust Flame 2023;248:112581. http://dx.doi.org/10.1016/j.combustflame.2022.112581.

[19] Bowman CT. Control of combustion-generated nitrogen oxide emissions: Technology driven by regulation. Symp (International) Combust 1992;24(1):859–78. http://dx.doi.org/10.1016/s0082-0784(06)80104-9.

[20] WESTENBERG AA. Kinetics of NO and CO in lean, premixed hydrocarbon-air flames. Combust Sci Technol 1971;4(1):59–64. http://dx.doi.org/10.1080/00102207108952472.

[21] Xi Z, Fu Z, Hu X, Sabir SW, Jiang Y. An experimental investigation on flame pulsation for a swirl non-premixed combustion. Energies 2018;11(7):1757. http://dx.doi.org/10.3390/en11071757.

[22] Valera-Medina A, Syred N, Bowen P, Crayford A. Studies of swirl burner characteristics, flame lengths and relative pressure amplitudes. J Fluids Eng 2011;133(10). http://dx.doi.org/10.1115/1.4004908.

[23] KIM H, ARGHODE V, GUPTA A. Flame characteristics of hydrogen-enriched methane–air premixed swirling flames. Int J Hydrog Energy 2009;34(2):1063–73. http://dx.doi.org/10.1016/j.ijhydene.2008.10.035.

[24] Kroniger D. Vorhersage der NO_x Emissionen einer wasserstoffbetriebenen Industriegasturbinenbrennkammer mit Wassereinspritzung [Ph.D. thesis], Rheinisch-Westfälische Technische Hochschule Aachen, RWTH Aachen University; 2019, p. 2019. http://dx.doi.org/10.18154/RWTH-2019-02596, URL http://publications.rwth-aachen.de/record/756624.

[25] Reichel TG, Terhaar S, Paschereit CO. Flashback resistance and fuel–air mixing in lean premixed hydrogen combustion. J Propuls Power 2018;34(3):690–701. http://dx.doi.org/10.2514/1.b36646.

[26] Vignat G, Durox D, Candel S. The suitability of different swirl number definitions for describing swirl flows: Accurate, common and (over-) simplified formulations. Prog Energy Combust Sci 2022;89:100969. http://dx.doi.org/10.1016/j.pecs.2021.100969.

[27] Marragou S, Magnes H, Poinsot T, Selle L, Schuller T. Stabilization regimes and pollutant emissions from a dual fuel CH4/H2 and dual swirl low NOx burner. Int J Hydrog Energy 2022;47(44):19275–88. http://dx.doi.org/10.1016/j.ijhydene.2022.04.033.

[28] Aswani, Praveen VV, Thangavelu S. Performance analysis of variants of differential evolution on multi-objective optimization problems. Indian J Sci Technol 2015;8(17). http://dx.doi.org/10.17485/ijst/2015/v8i17/65727.

[29] Chandrasekaran N, Guha A. Study of prediction methods for NOx emission from turbofan engines. J Propuls Power 2012;28(1):170–80. http://dx.doi.org/10.2514/1.b34245.

[30] TSAGUE L, TSOGO J, TATIETSE T. Prediction of the production of nitrogen oxide (NOx)(NOx) in turbojet engines. Atmos Environ 2006;40(29):5727–33. http://dx.doi.org/10.1016/j.atmosenv.2006.05.055.

[31] Tsague L, Tatietse TT, Ngundam J, Tsogo J. Prediction of emissions in turbojet engines exhausts: relationship between nitrogen oxides emission index (EINOx) and the operational parameters. Aerosp Sci Technol 2007;11(6):459–63. http://dx.doi.org/10.1016/j.ast.2007.04.005.

[32] Doost A, Ries F, Becker L, Bürkle S, Wagner S, Ebert V, Dreizler A, di Mare F, Sadiki A, Janicka J. Residence time calculations for complex swirling flow in a combustion chamber using large-eddy simulations. Chem Eng Sci 2016;156:97–114. http://dx.doi.org/10.1016/j.ces.2016.09.001.

[33] Pedersen LS, Breithauptb P, Johansen KD, Weber R. Residence time distributions in confined swirling flames. Combust Sci Technol 1997;127(1–6):251–73. http://dx.doi.org/10.1080/00102209708935696.

[34] Bürkle S, Becker LG, Agizza MA, Dreizler A, Wagner S. Comparison of two measurement strategies to obtain the residence time distribution in combustion chambers using tunable diode laser absorption spectroscopy. Appl Opt 2019;58(10):C36. http://dx.doi.org/10.1364/ao.58.000c36.

[35] Røkke N, Hustad J, Sønju O, Williams F. Scaling of nitric oxide emissions from buoyancy-dominated hydrocarbon turbulent-jet diffusion flames. Symp (International) Combust 1992;24(1):385–93. http://dx.doi.org/10.1016/s0082-0784(06)80050-0.

[36] WANG Z, LIN Y, WANG J, ZHANG C, PENG Z. Experimental study on NO emission correlation of fuel staged combustion in a LPP combustor at high pressure based on NO-chemiluminescence. Chin J Aeronaut 2020;33(2):550–60. http://dx.doi.org/10.1016/j.cja.2019.09.004.

[37] Xu S, Xi L, Tian S, Tu Y, Chen S, Zhang S, Liu H. Numerical investigation of pressure and H2O dilution effects on no formation and reduction pathways in pure hydrogen MILD combustion. Appl Energy 2023;350:121736. http://dx.doi.org/10.1016/j.apenergy.2023.121736.

[38] Lee J, Barbet MC, Meng Q, Cornell RE, Burke MP. Experimental support for a new NOx formation route via an hnno intermediate. Combust Flame 2023;257:112632. http://dx.doi.org/10.1016/j.combustflame.2023.112632.

[39] Durocher A, Bourque G, Bergthorson JM. Quantifying the effect of kinetic uncertainties on NO predictions at engine-relevant pressures in premixed methane–air flames. J Eng Gas Turbines Power 2020;142(6). http://dx.doi.org/10.1115/1.4047108.

[40] Choi J, Ahn M, Kwak S, Lee JG, Yoon Y. Flame structure and NOx emission characteristics in a single hydrogen combustor. Int J Hydrog Energy 2022;47(68):29542–53. http://dx.doi.org/10.1016/j.ijhydene.2022.06.247.

[41] Kuo KK-Y. Principles of combustion. 2nd ed. Nashville, TN: John Wiley & Sons; 2005.

[42] Sweeney MS, Hochgreb S, Dunn MJ, Barlow RS. The structure of turbulent stratified and premixed methane/air flames II: Swirling flows. Combust Flame 2012;159(9):2912–29. http://dx.doi.org/10.1016/j.combustflame.2012.05.014.

[43] Mecklin CJ, Mundfrom DJ. A Monte Carlo comparison of the type I and type II error rates of tests of multivariate normality. J Stat Comput Simul 2005;75(2):93–107. http://dx.doi.org/10.1080/0094965042000193233.