



Gas turbine emissions modeling: from empirical correlations to hybrid physics-machine learning frameworks

Tariq R. Aboalhol^{a,*}, Khaled A. Abdusamad^b

^a Libyan Higher Technical Center for Training and Production, Janzour, Libya

^b Mechanical Eng. & Energies Dept. - Libyan Academy for Postgraduate Studies, Janzour, Libya

ARTICLE INFO

Keywords:

Gas turbine emissions modeling
Nitrogen oxides
Carbon monoxide
Computational fluid dynamics
Chemical reactor networks
Machine learning
Large eddy simulation
Hybrid physics-machine learning models

ABSTRACT

This review examines 56 peer-reviewed studies on gas-turbine emission modeling, covering data-driven, semi-empirical, physics-based, and hybrid approaches. It compares model fidelity, computational cost, and interpretability across these methods and finds a clear hierarchy in which physics-based models provide the mechanistic foundation, while hybrid and data-driven systems extend that fidelity for scalable, real-time prediction. Together they establish a coherent basis for modeling NO_x and CO under both conventional and fuel-flexible conditions.

A structured search of Scopus, Web of Science, and supplementary databases (records = 260; screened = 230; full-text = 110; included = 56) followed PRISMA methodology to ensure transparency and reproducibility. Eligible studies offered quantitative validation of combustion or emission models and were classified into four methodological groups. Each was examined for governing variables, validation approach, and performance metrics, allowing comparison of predictive reliability across operating regimes.

Recent research shows that hybrid architectures, which couple mechanistic transparency with data-driven adaptability, are now dominant. Physics-based simulations generate the high-fidelity datasets used for training, whereas machine-learning surrogates interpolate rapidly within validated domains. Remaining challenges involve transient validation, fuel-flexible kinetics, limited domain coverage, and the absence of standardized uncertainty quantification. Adaptive, physics-guided learning frameworks that preserve interpretability and quantify reliability across fuels, loads, and stages of engine aging appear most capable of addressing these gaps.

Introduction

Gas turbines remain central to modern power generation and aviation. Reducing nitrogen oxide (NO_x) and carbon monoxide (CO) emissions has become a priority for both regulators and engineers. Recent policy developments, including ICAO Annex 16 amendments and national net-zero commitments, have intensified pressure on turbine manufacturers to sustain low emissions across variable load conditions and to operate effectively with emerging fuels such as hydrogen and sustainable aviation fuels. These fuels alter combustion behavior and increase the need for reliable emission models.

A broad range of modeling approaches has been developed to meet this challenge. Physics-based frameworks such as Reynolds-averaged Navier–Stokes (RANS), large-eddy simulation (LES), and chemical reactor networks (CRNs) offer detailed mechanistic insight but remain

computationally demanding. Semi-empirical and data-driven models provide faster predictions but require calibration within specific operating envelopes. Recent work, including Huang et al. [1] and Mubashir et al. [2], shows that hybrid frameworks integrating physical constraints with machine learning can improve generalization across operating regimes. This diversity reflects a critical operational conflict: regulatory frameworks increasingly require near-real-time emissions estimation across variable load and fuel conditions for compliance and monitoring, whereas designers require high-fidelity resolution to validate hydrogen and sustainable aviation fuels. No single paradigm satisfies both demands. As a result, cross-paradigm hybrid frameworks have emerged as a practical necessity rather than a methodological preference, combining mechanistic fidelity with the computational tractability required for regulatory compliance. However, existing reviews primarily address either combustion fundamentals or individual modeling

* Corresponding author.

E-mail address: taboalhol@gmail.com (T.R. Aboalhol).

<https://doi.org/10.1016/j.nexres.2026.101332>

Received 10 July 2025; Received in revised form 29 December 2025; Accepted 9 January 2026

Available online 9 January 2026

3050-4759/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

classes, without systematically comparing predictive accuracy, computational efficiency, and regulatory relevance across paradigms for low-emission turbine concepts.

We fill these gaps through a comparative analysis of NO_x and CO emission models spanning four categories: data-driven, semi-empirical, physics-based, and hybrid frameworks. Emphasis is placed on performance under realistic operational scenarios such as part-load transients and fuel-flexible combustion.

To address this methodological and regulatory disconnect, we present a comparative analysis of NO_x and CO emission models spanning four categories: data-driven, semi-empirical, physics-based, and hybrid frameworks. Emphasis is placed on performance under realistic operational scenarios such as part-load transients and fuel-flexible combustion.

This work provides the first integrated benchmarking of hybrid, data-driven, and physics-based gas-turbine emission models using standardized metrics and validation criteria.

Review methodology

This review follows the narrative-synthesis framework of Ferrari [3] adapted for gas-turbine emission modeling. The workflow in Fig. 1 aligns with PRISMA principles. Inclusion and exclusion criteria are listed in Table 1 for transparency.

Search strategy

The objective was to identify quantitative modeling studies of NO_x and CO emissions in gas turbines, including empirical, physics-based (CFD and CRN), and data-driven (machine-learning and hybrid) frameworks. Searches spanning January 2000 to September 2025 used structured keyword queries across Scopus (primary) and Web of Science (secondary). Representative queries included:

- “gas turbine” AND “NO_x emission model”

Table 1

Inclusion and exclusion criteria for study selection.

Criterion	Inclusion	Exclusion
Publication Type	Peer-reviewed journal or conference papers	Review papers; non-peer-reviewed sources
Language	English	Non-English publications
Time Frame	January 2000 - September 2025	Outside this period
Model Focus	Empirical, physics-based (CFD/CRN), ML, or hybrid models	Purely experimental or measurement studies without modeling
Validation	Quantitative validation via experiments or detailed simulations	No reported performance metrics or validation data
Accessibility	Full-text available	Abstract-only or inaccessible full text

These criteria ensured methodological consistency across 56 studies and enabled structured synthesis and quantitative comparison.

- “gas turbine” AND “CO prediction”
- (“CFD” OR “reactor network” OR “machine learning”) AND “gas turbine emissions”

A total of 260 records were retrieved. After duplicate removal ($n = 230$), titles and abstracts were screened. One-hundred-ten full-text articles were assessed, and 56 studies met all inclusion criteria for narrative synthesis.

Inclusion and exclusion criteria

Following Ferrari [3] and PRISMA guidelines, studies were included only if they presented quantitative, validated emission models. Eligible publications developed or evaluated NO_x or CO prediction models for gas turbines using experimental data or high-fidelity simulations, such as LES or CRN–CFD coupling. All included studies were peer-reviewed, published in English, and available in full text.

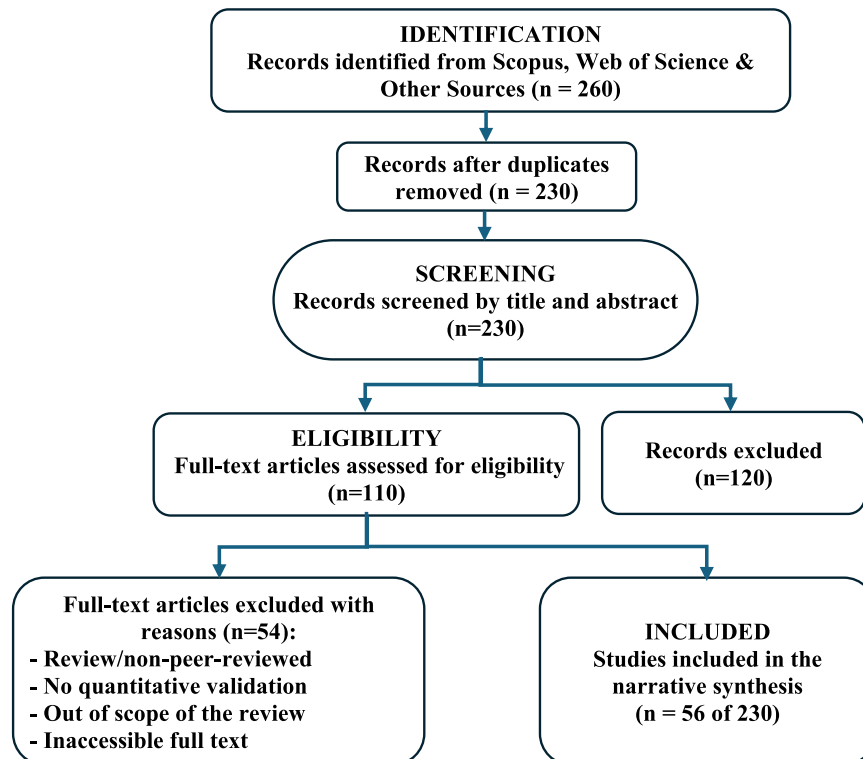


Fig. 1. PRISMA flow diagram illustrating the study selection process.

Screening and selection

After duplicates were removed, 230 records were screened by title and abstract. Studies unrelated to gas-turbine emissions, lacking modeling content, or without quantitative validation were excluded. One-hundred-ten full-text papers were evaluated against the criteria listed in Table 1, of which fifty-six satisfied all requirements and were included in the final synthesis. To ensure screening consistency and methodological reproducibility, title–abstract and full-text screening were conducted independently by two reviewers using the same inclusion criteria. Discrepancies were resolved through consensus discussion, and all final exclusion decisions were documented prior to synthesis. Data extraction followed a shared template applied independently and cross-checked before consolidation. This dual-reviewer process reduces selection bias and enhances transparency in accordance with PRISMA recommendations. The workflow is summarized in Fig. 1.

Data extraction and synthesis

Data extraction followed a standardized template to ensure comparability across modeling paradigms. For each study, the following were recorded:

1. Model category (empirical, CFD, CRN, ML, or hybrid)
2. Target species (NO_x , CO, or both)
3. Input and boundary parameters (e.g., inlet temperature, fuel flow, load condition)
4. Validation type (test rig, engine data, or benchmark simulation)
5. Performance metrics (RMSE, MAE, R^2 , or other quantitative indicators)
6. Reported strengths and limitations (as stated or inferred)

These data formed the basis for identifying methodological trends, assessing predictive performance, and evaluating generalization limits. Sections 3 and 4 present the comparative analyses derived from this dataset.

Emissions modeling approaches

Gas-turbine emission modeling has evolved from classical semi-empirical correlations to modern data-driven regressions that aim to balance physical interpretability with computational efficiency. These models typically use engine and ambient conditions such as inlet temperature, pressure ratio, fuel-air ratio, and load to estimate NO_x and CO emissions with minimal computational cost. CO_2 is usually derived from the fuel's carbon balance rather than empirical fitting. Such models are used in design screening and online monitoring, often coupled with CFD-based reactor networks for added physical consistency [4].

Data-Driven and semi-empirical models

CFD provides detailed insight into combustion and emissions but remains too computationally expensive for continuous design, monitoring, or control. Semi-empirical and data-driven models estimate NO_x and CO from measurable variables using correlations derived from experiments, field data, or CFD outputs. They run rapidly, but accuracy hinges on how well the training domain is represented. Reliability improves when inputs encode the physical drivers of pollutant formation such as temperature, residence time, and equivalence ratio.

Early semi-empirical work established predictive relationships between operating conditions and emissions. Chandrasekaran and Guha [5] linked inlet pressure, temperature, and humidity to NO_x using ICAO LTO data. Xi et al. [6] extended this approach with power-law fits incorporating flow rates and derived flame parameters. While both models captured the dominant trends, they assumed steady operation and fixed fuel composition, which limits their transferability to transient

conditions or alternative fuels due to quasi-steady chemistry assumptions and the absence of explicit links between residence time and burnout.

As transient and variable-load prediction gained importance, researchers turned to operational datasets. Katzenstein and Apt [7] analyzed minute-resolution turbine data and identified emission penalties during load ramping. Kochueva and Nikolskii [8] applied symbolic regression with fuzzy clustering to 36 000 samples, achieving high accuracy for CO and NO_x and demonstrating that calibration on real-engine data improves reliability. However, in the absence of explicit physical constraints, extrapolation beyond the calibration envelope remains fragile.

Recent studies integrate physical descriptors into statistical models. Zhao et al. [9,10] and Bajwa et al. [11] showed that embedding mechanistic variables improves generalization while preserving computational speed, reporting MAPE of approximately 5.3 % and MAE of approximately 5.6 % under transient loads when features represent kinetic and mixing scales.

Regression remains central for plant-scale monitoring. Egware and Kwasi-Effah [12] obtained very high internal R^2 values for CO_2 over four years, showing that strong correlation within a narrow dataset does not ensure validity under new fuels or off-design conditions. In the absence of uncertainty quantification, overfitting risk remains hidden. Domain checks, prediction intervals, and confidence bands are therefore required to prevent narrow-range fits from being misinterpreted as general reliability.

Table 2 traces the shift from curve-fits to CFD-assisted surrogates. Accuracy depends less on algebraic form than on how clearly a model represents pollutant-formation physics. Efficiency and fidelity are not in conflict; generalization improves when learned relationships retain transparent links to combustion chemistry and flow behavior.

Physics-Based emission models

Physics-based models estimate NO_x and CO by directly solving the governing equations for flow, turbulence, and chemical kinetics. Unlike semi-empirical or data-driven methods, they capture the mixing and reaction processes that control pollutant formation. These models achieve high predictive accuracy but remain computationally intensive and highly sensitive to boundary conditions, turbulence closure, and kinetic detail.

These approaches include reduced-chemistry CFD, which employs skeletal or truncated mechanisms to make three-dimensional simulations feasible, and chemical reactor networks (CRNs), which model combustion as interconnected ideal reactors that capture dominant kinetics at lower computational cost [1,13].

Early frameworks coupled CFD and CRNs to balance fidelity and efficiency. Novosselov and Malte [14] combined RANS-based CFD with an eight-step methane mechanism, reproducing NO_x and CO trends in laboratory and industrial combustors. Park et al. [15] arranged perfectly stirred and plug-flow reactors informed by CFD, matching test-rig data within experimental uncertainty while reducing runtime to the order of minutes.

As computing power increased, fully resolved CFD became more common. Filippov et al. [16] applied SST-RANS with detailed NO_x kinetics, achieving a normalized RMS deviation of approximately 0.15 in bench tests. Innocenti et al. [17] extended the analysis to aero-engine conditions, showing good agreement overall but up to 25 % CO error at idle due to localized quenching. These discrepancies confirm that predictive accuracy degrades when flame structure and temperature gradients are insufficiently resolved.

The flamelet-generated manifold (FGM) method, introduced by van Oijen and de Goey [18], couples detailed flamelet chemistry with manifold reduction to efficiently represent multidimensional premixed flames. Goldin et al. [19] showed that unsteady flamelet formulations capture transient CO and OH evolution more accurately than steady

Table 2

Data-driven and semi-empirical emissions models.

Study	Model Type	Inputs	Targets	Validation Data	Validation Context	Performance
Chandrasekaran & Guha [5]	Semi-empirical correlation	Inlet P, T, humidity	NOx	ICAO LTO database	6 turbofan models, LTO cycle	Good agreement vs. P3-T3 & ICAO data; no metrics
Xi et al. [6]	Empirical power-law fit	Inlet P; CH ₄ & air flow; derived flame params.	NOx, CO	12 steady cases (3–6 bar)	20 kW swirl combustor with C-D nozzle and gas probe	Good agreement; exponents reported
Zhao et al. [9]	Semi-empirical regression (CFD-based)	Inlet P, air flow, CFD-derived T _{pz}	NOx	CFD dataset (~5.9 M cells)	Single-flame-tube CFD; parametric runs	No R ² or RMSE reported
Kochueva & Nikolskii [8]	Symbolic regression + fuzzy clustering	Turbine operating params.	NOx, CO	Industrial field data	36 733 hourly samples (5 yr)	R ² ≈ 0.89 (CO); R ² ≈ 0.83 (NOx)
Egware & Kwasi-Effah [12]	Multivariate linear regression	Inlet T, humidity, P-ratio, exhaust flow	CO ₂	Ebsilon simulations	Calibrated with 4 yr operational data	R ² = 0.998
Zhao et al. [10]	CFD-derived polynomial surrogate	6 inlet boundary conditions	NOx	CFD-generated data	152 OSF design points	MAPE = 5.34 %
Bajwa et al. [11]	Semi-empirical (physics-based)	Cycle-resolved pressure	NOx	Test rig (HSDI diesel engine)	Transient load steps	MAE = 5.57 %; R ² = 0.70

Notes: Concentrations are reported as in the original sources. For reference, at 25 °C and 1 atm, $\text{mg m}^{-3} = (\text{ppmv} \times \text{MW}) / 24.45$, where MW is molecular weight (g mol^{-1}). The NO_x-CO trade-off reflects a core combustion constraint targeted by these models.

variants. Klarmann et al. [20,21] incorporated stretch and heat-loss effects, improving CO predictions at part-load conditions. Perpignan et al. [22] applied FGM to an inter-turbine burner and reproduced NO_x trends, although CO was overpredicted due to incomplete quenching treatment. More recently, Wang et al. [23] and Tenango-Pirín et al. [24] reported prediction errors below 4 % for both pollutants using unsteady FGM in micro- and laboratory-scale combustors.

Thickened-flame models (TFM) and artificially thickened flame

models (ATFM) remain practical tools for LES of premixed combustion. Castellani et al. [25] validated flame-structure predictions in gas-turbine geometries, illustrating current practice. Finite-rate chemistry has also been implemented in RANS and LES through PaSR, EDC, and EDM/FRC closures to resolve detailed kinetics and pollutant pathways [26–28]. These explicit rate-based models improve temperature sensitivity for NO_x and CO predictions under lean-premixed operation.

Recent work emphasizes validation at elevated pressure. Ding et al.

Table 3

Physics-Based Emission Models.

Study	Modeling Approach	Target Species	Validation Platform	Validation Cases (n)	Press/Temp Range	Runtime	Key Findings / Accuracy
Novoselov & Malte [14]	RANS-CFD + CRN (8-step CH ₄ mech.)	NOx, CO	Bench & industrial combustors	8 + 5	5–20 atm; tests ≈14 atm	Not reported	Good agreement between CFD and CRN predictions
Wegner et al. [32]	CFD (RANS, two-timescale model TSS)	CO (primary); UHC qualitative	Part-load burner rig	3	1 bar (part-load)	N/A	Captured CO burnout via second progress variable; validated vs lab rig
Goldin et al. [19]	Flamelet-Generated Manifolds (1D/0D)	CO, OH	Computational flamelet sim. (no exp.)	Multiple	1 atm; 300 K inlet	N/A	CO and OH evolution sensitive to steady vs unsteady flamelets
Park et al. [15]	CFD-guided CRN (PSR + PFR)	NOx, CO	Industrial DGT5 rig	13	1 bar; 305–417 °C inlet	Few CPU min (CRN)	Reproduced NOx & CO within exp. uncertainty
Andreini et al. [33]	Eulerian-Lagrangian CFD (film + spray)	- (fuel-spray only)	Aero-engine PERM injector rig	3 (Idle-Full)	5.3–22.4 bar; 613–811 K	Not reported	Resolved atomization / film dynamics; no emission stage
Klarmann et al. [20]	FGM w/ stretch & heat-loss correction	-	Atmospheric burner rig	1	1 bar; ~673 K inlet	N/A	Improved flame-brush prediction without extra dims
Filippov et al. [16]	SST-RANS + detailed NOx kinetics	NOx	Bench-scale combustor	4	1 bar; exhaust ≈ 1500 °C	Not reported	Normalized RMS deviation ≈ 0.15 (steady-state comparison)
Innocenti et al. [17]	k- ω SST CFD → CRN + diffusion-flux	NOx, CO	Aero-engine PERM rig	3 (Idle-Cruise-Takeoff)	5.3–22.4 bar; 613–811 K	Few CPU min (CRN)	NOx error ≈ 20 %; CO ≈ 26 % at idle (extinction)
Klarmann et al. [21]	Extended FGM + CO transport + quenching	CO	Part-load burner rig	5	1 bar; 300 °C preheat	Not reported	Matched measured CO; outperformed single-variable FGM
Perpignan et al. [22]	RANS (k- ϵ) + FGM + NOx transport	NOx, CO	18° sector ITB rig	18	1 bar; preheated air	Not reported	Captured NOx trend; overpredicted CO; missed non-monotonicity
Fooladgar et al. [26]	CFD (RANS) + Finite-Rate Chem. (EDC-type)	Radicals, CO, soot precursors	Flat-flame burner	Qualitative	1 atm	Not reported	Qualitative agreement for radicals & T; limited quant. fit
Ding et al. [29]	CFD (RANS) + FGM	NOx (main)	F-class engine data	2	~16 MPa; 1300–1500 K	Not reported	Rel. errors +1.42 % and –5.5 % vs data
Tenango-Pirín et al. [24]	CFD (RANS / Unsteady Flamelet)	NOx, CO	Published exp. flame data	3 (unsteady)	Not stated	Not reported	Low avg error CO ≈ 2.2 % (steady validation)
Wang et al. [23]	CFD (RANS / Unsteady Flamelet)	NOx	Micro-combustor rig	8 heat loads	0.36 MPa	Not reported	Max local rel. error 3.56 %
Cheng et al. [34]	Physics-based 1D CFD + Exp.	NOx, CO	Counterflow burner rig	Few (limited)	1 atm	Not reported	Qualitative agreement (laminar / turbulent study)
Hajivand [27]	CFD (RANS) + Finite-Rate Chem. (EDM/FRC)	NOx	TAY combustor rig	9 model variants	1 atm	Not reported	Best (SST + EDM/FRC) mean deviation ≈ 2.3 % NOx
Huang et al. [1]	3D RANS + Finite-Rate Chem. + NOx post-proc.	NOx	Industrial turbine rig	5	1.5–2.3 MPa; 650–780 K	Not reported	NOx error ± 4.2 % vs measured

[29] simulated startup of an F-class DLN 2.0+ turbine using RANS-FGM, matching field data within $\pm 5\%$ at pressures of 0.9–1.5 MPa. Hajivand [27] compared nine turbulence–chemistry interaction models in a TAY rig and found the SST combined with EDM/FRC closure to be the most accurate, with a mean deviation of approximately 2.3 %. Huang et al. [1] applied finite-rate RANS to combustors operating at 1.5–2.3 MPa and achieved errors within $\pm 4.2\%$ relative to measurements, confirming increasing model maturity under industrial conditions.

Finite-rate and hybrid LES–CRN frameworks also clarify reaction pathways. Beige and Mardani [30] used LES–CRN coupling to isolate dominant NO formation loops, while Starik et al. [31] showed that kinetic-mechanism selection alone can alter predicted NO_x levels by $>50\%$.

Despite major advances, physics-based models still require detailed kinetic data and accurate boundary specification, limiting their use for transients and fuel-flexible operation. Their predictive strength depends on turbulence-chemistry coupling and on resolving slow CO oxidation and quenching near cold walls. Evidence indicates that reliability scales with the fidelity of coupled transport and chemistry timescales rather than model size. Unsteady FGM and finite-rate RANS or LES achieve higher accuracy when both fast-flame and slow-burnout regimes are resolved. The main challenge is to retain mechanistic detail while reducing cost, now pursued through hybrid and physics-informed machine-learning frameworks [13]. Table 3.

Machine-Learning and hybrid frameworks

Machine-learning methods complement physics-based simulations, which are limited in handling transient behavior, fuel variability, and component aging. These models learn nonlinear patterns from operational data to allow near-real-time prediction for monitoring and control. Hybrid architectures are designed to couple first-principles solver outputs or constraints with statistical learning methods such as CFD-trained surrogates, CRN-informed models, and physics-constrained Bayesian networks. Their reliability depends on representative sampling, disciplined temporal partitioning, and physically meaningful features [35,36].

Early work refined semi-empirical correlations rather than replacing them. Lipperheide et al. [35] embedded degradation corrections in a predictive emissions monitoring system (PEMS) workflow, reducing NO_x error relative to baseline correlations. Using the UCI gas-turbine emissions dataset (36 733 hourly samples), Kaya et al. [36] showed that maintaining chronological order during validation prevents inflated accuracy. Si et al. [37] achieved U.S. EPA-level precision for NO_x through disciplined tuning of network depth and learning rate, demonstrating that data management is as critical as network architecture.

Later studies sought accuracy with interpretability. Kochueva and Nikolskii [8] used symbolic regression with fuzzy clustering to derive compact analytical forms for NO_x and CO ($R^2 \approx 0.83$ – 0.89). Niño-Adán et al. [38] examined normalization effects on convergence and feature attribution. Dirik [39] coupled a neuro-fuzzy model with genetic optimization to capture nonlinear combustion behavior, while Li et al. [40] embedded neural networks within multi-objective optimization linking emissions to engine-performance trade-offs.

After 2023, focus shifted toward robustness and fuel flexibility. Potts et al. [41] showed that gradient-boosted ensembles can match reduced-kinetic PEMS accuracy with systematic feature selection. Aslan [42] and Hoque et al. [43] achieved $R^2 > 0.92$ using tuned ensembles on standardized datasets, marking the move from exploratory modeling to reproducible benchmarking. Modern pipelines employ nested cross-validation, rolling-window retraining, and Bayesian or genetic tuning with variance-aware temporal segmentation, improving stability under long-term drift.

Recent advances integrate ML flexibility with physical constraints. Reumschüssel et al. [44] trained Bayesian surrogates on swirl-rig

experiments to interpolate NO_x responses while quantifying uncertainty. Zhao et al. [10] built a polynomial surrogate from optimally sampled CFD data (MAPE $\approx 5\%$). Mubashir et al. [2] trained support-vector models on CFD-generated datasets, achieving narrow confidence intervals and fast parametric exploration. These physics-guided surrogates retain explicit variables such as flame temperature and residence time, enabling extrapolation across fuels and loads with mechanistic transparency.

The key limitation remains fragile generalization. Models trained on narrow or static datasets lose accuracy under distributional drift. Weak validation discipline causes data leakage or inflated metrics [45]. Some studies rely on proxy targets lacking physical meaning, while unconstrained networks may violate conservation laws unless regularized with physics-aware features. Data scope and constraint formulation, not algorithm depth, determine long-term reliability.

Across the literature, progress follows a consistent trajectory from empirical correlation to constraint-embedded learning. The defining advance is integrating physical knowledge directly into data pipelines, uniting interpretability with adaptability. This evolution positions ML as a quantitative complement to physics, forming the foundation for the hybrid architectures examined in Section 3.4. Further benchmarking using various ensemble and recurrent architectures on both standardized and proprietary engine datasets provides a broader basis for model comparison [46–51]. Table 4.

Advanced physics-based extensions

Advanced physics-based frameworks extend conventional CFD and LES to capture turbulence-chemistry coupling, pressure-dependent kinetics, and wall heat loss, all of which govern NO_x and CO formation in modern combustors. These models resolve mechanisms controlling emissions under realistic, fuel-flexible conditions.

LES remains the leading approach as it resolves the turbulent structures that drive mixing and flame stabilization while maintaining feasible computational cost. Yunoki et al. [52] showed that non-adiabatic LES captures wall heat loss causing CO buildup under lean operation. Chikishev et al. [53] used Proper Orthogonal Decomposition to link coherent vortical modes to NO and CO fluctuations, revealing structure-emission coupling. Castellani et al. [25] applied an artificially thickened-flame LES to hydrogen-methane blends, reproducing flame topology and radical pathways. These studies confirm that LES recovers transient emission physics, although accuracy depends on subgrid closures and kinetic mechanisms.

CFD–CRN coupling enhances chemical resolution by feeding reactor networks with residence-time and scalar data from CFD. Fichet et al. [4] and Starik et al. [31] showed that predicted NO_x can vary by over 50 % with kinetic-mechanism choice, exposing sensitivities absent in global models. Beige and Mardani [30] coupled LES with CRNs to isolate radical-formation zones and quantify equivalence-ratio effects on NO yield. This selective coupling preserves mechanistic clarity while moderating computational cost, identifying pathways that guide design optimization.

Physics-informed surrogates accelerate parametric analysis. Zhao et al. [10] built a polynomial surrogate from space-filling CFD datasets (MAPE $\approx 5\%$). Mubashir et al. [2] trained support-vector regressors on CFD-generated data to predict NO_x with ppm-level precision, enabling efficient burner optimization. These reduced-order models retain physical relevance through features anchored to combustion-governing variables.

This paradigm marks a shift from reproducing averaged flow behavior to resolving localized events that control pollutant formation. Advanced simulations now capture the transient multiscale structures that supply the mechanistic data and physical constraints feeding reduced-order and hybrid models. High-fidelity physics provides the empirical foundation for training and bounding tractable frameworks. Predictive accuracy still depends on turbulence-chemistry fidelity, and

Table 4

Representative Machine-Learning and Hybrid Frameworks for Gas-Turbine Emission Prediction.

Study (year)	Model (short)	Targets	Data source & size	Train / val / test split	Key performance (primary metric)	Notes
Lipperheide et al. [35]	Semi-empirical PEMS with aging corrections	NOx	Three heavy-duty turbines; ~370 000 points (3 yr)	Commissioning for formulation; 3-yr ops test	RMSE reduced 54 % to 1.26 ppm	Aging-calibrated correlation vs baseline
Kaya et al. [36]	ELM	CO, NOx	UCI PEMS (2011–2015), 36 733	Chronological 2011–12 / 2013 / 2014–15	MAE CO = 0.89 mg m ⁻³ ; MAE NOx = 7.27 mg m ⁻³	Temporal order preserved
Si et al. [37]	ANN (Keras)	NOx	CEMS, 15 106 hourly	80/20; internal 90/10 val	MAE = 0.5982; $r = 0.9451$; diff = 0.14 %	Open-source Keras; 32–64–64–64; Nadam
Kochueva & Nikolskii [8] †	Symbolic regression + fuzzy clustering	NOx, CO	UCI PEMS ~37 000	70/30 on subsets	$R^2 \approx 0.89$ (CO); $R^2 \approx 0.83$ (NOx)	Compact symbolic forms; cluster-dependent
Li et al. [40]	ANN for optimization	NOx (+others)	F-class engine ops (proprietary)	70/15/15	$R^2 \approx 0.96$ (NOx)	Combustion/ops tuning
Niño-Adán et al. [38]	ANN + normalization study	NOx, CO	UCI PEMS ~37 000	70/30	Normalization improved convergence/attribution	Feature relevance examined
Dirik [39]	ANFIS + GA	NOx	UCI PEMS ~37 000	70/30	$R = 0.899$; RMSE = 5.16 mg m ⁻³	GA-optimized neuro-fuzzy
Aquize et al. [46]	NARX (recurrent)	NOx (context)	Siemens SGT6–5000F ops data; 2 305 records	70/15/15	RMSE = 0.4411 % (relative); $R^2 \approx 1$	Internal dataset; temporal forecasting
Faqih et al. [47]	XGBoost + K-Means	NOx, CO	DLE engine ops; ~100 000	70/30	$R^2(\text{NO}_x) = 0.93$; $R^2(\text{CO}) = 0.71$	Semi-supervised segmentation
Potts et al. [41]	XGBoost (vs kinetics PEMS)	NOx	Siemens SGT400: full = 5860×174 ; cropped = 6429×88	10-fold CV	RMSE full = 4.08 ppm; cropped = 0.95 ppm	Cropped feature set more stable
Aslan [42]	Ensembles (various)	NOx	UCI PEMS 36 733	80/20	$R^2 = 0.9204$; RMSE = 3.95 ppm; MAE = 2.17 ppm	Tuned ensembles; benchmarked
Coelho et al. [48]	Deep Forest (ensemble)	CO, NOx	UCI PEMS 36 733	5-fold CV	CO: $R^2 \approx 0.54$, RMSE = 1.45 mg m ⁻³ ; NOx: $R^2 \approx 0.47$, RMSE = 7.65 mg m ⁻³	Favored interpretability
Dai et al. [49]	TSECIT2FNN-LSTM	NOx	UCI PEMS ~37 000	60/40	$R^2 \approx 0.9997$	Self-evolving fuzzy rules
Hoque et al. [43]	Ensembles (benchmark)	NOx	UCI PEMS 36 733	80/20	$R^2 = 0.9276$; RMSE = 2.93 ppm; MAPE = 3.45 %	Broad model comparison
Reumschüssel et al. [44]	Bayesian surrogates (hybrid exp+ML)	NOx (+others)	Swirl rig; 76 points (NOx)	CV on exp. points	Rel. error = 4.4 % (NOx)	Design optimization + UQ
Shaik et al. [50]	XGBoost	Energy yield (proxy)	UCI PEMS ~36 000	80/20	$R^2 = 0.9983$; RMSE = 0.6412; MAE = 0.4616	Proxy target; limited interpretability
Mubashir et al. [2]	Hybrid CFD-ML (SVR surrogate)	NOx	CFD-generated dataset	5-fold CV	RMSE ≤ 0.0023 ; MAE ≤ 0.0023 ; CI ± 1.8 ppm	SVR with CIs; fast exploration
Sinap [51]	Stacking ensemble (GBDT/LGBM/CatBoost+XGB)	NOx, CO	UCI PEMS 36 733 hourly	80/20	NOx: RMSE = 2.77 mg m ⁻³ , $R^2 = 0.9194$; CO: RMSE = 0.46 mg m ⁻³ , $R^2 = 0.8556$	Outlier handling improved stability
Winarto et al. [45]	Stacked ensemble	NOx, CO (+others)	Engine PEMS (size n/r)	Not specified	CO: RMSE = 0.00959, MAE = 0.00439, MAPE = 0.00300, $R^2 = 1.0000$; NOx: RMSE = 0.00461, MAE = 0.00140, MAPE = 0.00642, $R^2 = 1.0000$	Dataset size/split not reported; interpret with caution

Notes. Concentrations are reported as in the original studies. At 25 °C and 1 atm, $\text{mg m}^{-3} \approx (\text{ppmv} \times \text{MW}) / 24.45$, where MW is the molecular weight (g mol^{-1}). Metrics shown correspond to each study's stated best-case test performance. Runtime information follows the cited publications; several omitted this parameter, underscoring the continued lack of standardized computational-cost reporting.

† Also listed in Table 2 as a transitional study using symbolic regression to derive an interpretable semi-empirical relation.

Studies are listed chronologically to illustrate the evolution of model architectures; citation numbers reflect the order of first appearance in the main text.

the main challenge is preserving this level of mechanistic detail at lower cost, which continues to drive the move toward hybrid and physics-informed methods. Table 5.

The progression from empirical correlations to advanced physics-based extensions forms a continuous expansion of mechanistic detail rather than a sequence of discrete replacements. Each paradigm carries forward characteristic constraints: empirical models are limited by data scope, physics-based models by computational cost, and machine-learning models by weak physical grounding. The prevailing trend across the literature is the emergence of hybrid architectures explicitly designed to overcome these inherited constraints. This evolution from separate methods to unified frameworks establishes the conceptual foundation for the synthesis developed in Section 4.

Cross-Method synthesis and discussion

This review examines four modeling paradigms: data-driven and semi-empirical, physics-based, machine-learning, and hybrid physics-

ML, within a unified comparative framework. Each balances predictive fidelity, interpretability, and computational cost differently. Three trade-offs emerge: speed versus fidelity, data requirements versus interpretability, and static versus adaptive frameworks.

Speed vs. fidelity

Low-fidelity models, including semi-empirical correlations and reduced-order CRNs, are optimized for computational speed, producing results within milliseconds to a few CPU minutes [6,15]. They are widely used for online monitoring, control, and rapid design screening where responsiveness is essential. Accuracy decreases under transient or fuel-flexible operation because simplified transport and kinetic assumptions cannot capture quenching, stratification, or unsteady behavior.

High-fidelity approaches such as RANS-based CFD and LES resolve turbulence-chemistry coupling and provide detailed information on flame stabilization and pollutant formation [1,25,55]. These models are

Table 5
Advanced Physics-Based Emission Modeling Frameworks.

Study	Approach	Target species	Validation	Validation cases / dataset	Operating conditions (highlights)	Key outcome / accuracy
Biagioli & Güthe [54]	LES with prompt / post-flame NOx (Damköhler split)	NOx	ALSTOM EV double-cone burner (exp + LES)	$n = 23$ (8 at 1 bar; 15 at 15 bar)	1 and 15 bar; varied ϕ	Unmixedness reversed the NOx pressure-exponent trend.
Fichet et al. [4]	Automated RANS $\rightarrow$ CRN mapping (PSRs)	NOx	Lab flame-tube (NG)	3 loads (100 %, 75 %, 50 %)	15 / 11.5 / 8.7 bar	NOx error ≈ 3 % nominal; ≈ 11 % at 50 % load.
Starik et al. [31]	CRN mechanism-sensitivity study	NOx	Diffusion and homogeneous combustors	3 mechanisms compared	7–12 bar; CH ₄ and kerosene	NOx predictions varied > 50 % with mechanism choice.
Serbin et al. [55]	3-D CFD (EDC + DPM; detailed / reduced kinetics)	NOx, CO	Dual-fuel low-emission combustor vs literature	4 cases	2–5 bar; $T_a \approx 770$ K	Predicted NOx ≈ 32 ppm, CO ≈ 9 ppm for radial injection.
Yunoki et al. [52]	LES (non-adiabatic FGM)	CO (+ qualitative NOx)	Literature test rig	Qualitative comparison	5 bar; $T_{in} \approx 673$ K; cooled walls	Heat loss increased CO at lean operation.
Chikishev et al. [53]	LES (FGM) + POD of coherent structures	NOx, CO	PIV-validated rig	$n = 3$ cases	≤ 3.4 atm; $T_{in} \approx 500$ K	POD linked coherent vortices to localized NOx/CO fluctuations.
Beige & Mardani [30]	Hybrid LES (LES + CRN coupling)	NOx	DLR GTMC rig (literature)	Qualitative comparison	Atmospheric	Identified vortical zones enhancing radical formation.
Bode et al. [56]	Physics-informed GAN sub-filter (PIESRGAN)	Combustion scalars	DNS of iso-octane flame kernel	DNS database (single kernel)	≈ 6 bar DNS	GAN reconstructed fine LES-scale fields consistently.
Castellani et al. [25]	LES with Artificially Thickened Flame (ATFM)	Flame dynamics (H ₂ blends)	Cardiff GTRC rig (OH*, flame shape)	Quantitative + qualitative tests	Atmospheric; H ₂ -CH ₄ blends	ATFM reproduced OH* and flame-shape trends with H ₂ fraction.
Zhao et al. [10]	CFD-ML surrogate (space-filling design)	NOx	CFD training dataset	152 design samples	Parametric range	MAPE = 5.34 %; validated CFD-ML workflow.
Ercole et al. [28]	LES comparison (PaSR vs EDC / FPV)	Flame dynamics	CECOST burner (PIV literature)	Comparative cases	Atmospheric / lab	Quantified model-dependent flame-structure differences.
Mubashir et al. [2]	Hybrid CFD-ML (SVR surrogate)	NOx	CFD dataset; 5-fold CV	Design space unspecified	Not reported	RMSE = 0.0023; CI ± 1.8 ppm; optimized burner design.
Zhou & Wang [57]	LES with Second-Order Moment (SOM) model	Flame dynamics	Literature swirling non-premixed flames	Qualitative comparison	Not specified	SOM captured macroscopic flame dynamics with qualitative agreement.

Notes: Concentrations are as reported in the original studies. At 25 °C and 1 atm, $\text{mg m}^{-3} = (\text{ppmv} \times \text{MW}) / 24.45$, where MW is molecular weight (g mol^{-1}). Operating conditions (pressure, temperature, ϕ) follow original sources; not specified indicates missing data. Runtime metrics are reported when available, reflecting the persistent absence of standardized computational-cost reporting.

effective for analyzing low-emission combustion and alternative fuels such as hydrogen and ammonia [28]. However, these simulations often require from tens to several thousand CPU hours per case, as reported by Filippov et al. [16] and Castellani et al. [25], limiting their practicality for large design sweeps or real-time use.

Hybrid strategies seek a balance between accuracy and efficiency by combining CFD or experimental data with reduced-order or ML

surrogates. Examples include CFD-informed CRNs [4,30], and surrogates trained on CFD-generated datasets [2,10]. These frameworks retain mechanistic fidelity while reducing runtime by one to two orders of magnitude, although their reliability outside trained regimes remains uncertain, particularly during transients.

A recurring limitation, illustrated in Fig. 2, is the inconsistent reporting of runtime. Many studies describe models as “low-cost” or

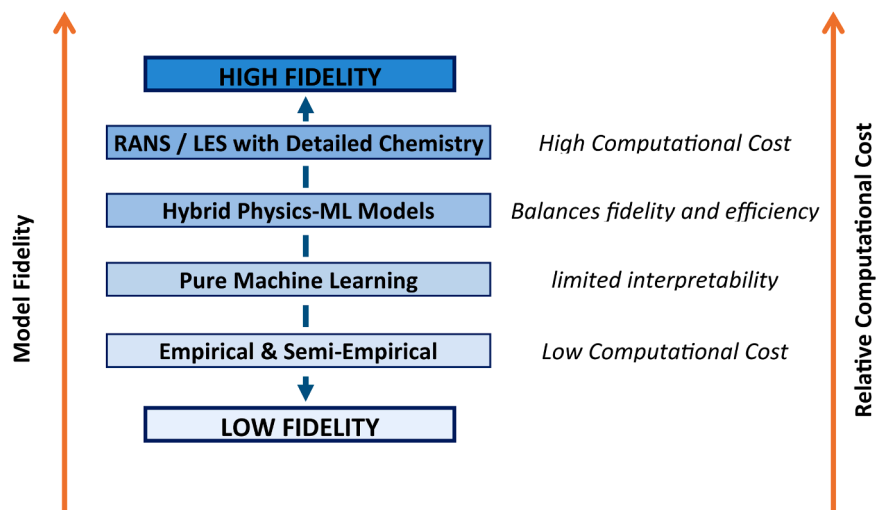


Fig. 2. Computational cost versus model fidelity.

Qualitative hierarchy of modeling approaches with indicative computational trends. Approximate runtime ranges adapted from Park et al. [15], Filippov et al. [16], and Castellani et al. [25].

“computationally intensive” without specifying solver settings, hardware, or parallelization method. This inconsistency hinders objective evaluation of computational trade-offs and obscures practical scalability. Standardized reporting of solver settings, hardware, and parallelization would markedly improve transparency and comparability. To contextualize the “Not Reported” entries, we provide an illustrative synthesis of computational cost based on representative configurations reported in the reviewed literature. For industrial-scale combustor sectors modeled with contemporary mesh resolutions, empirical and semi-empirical models typically execute in sub-second time on standard CPUs. Reduced-order chemical reactor network (CRN) models generally converge within seconds to minutes per operating point, depending on network size and chemistry reduction; for example, Fichet et al. [4] reported runtimes on the order of 10^3 seconds for networks comprising several hundred reactors.

In contrast, high-fidelity simulations impose substantially higher computational demands. Finite-rate RANS calculations commonly require on the order of 10^3 processor-hours per case [26], while recent LES studies employ meshes exceeding several million cells and incur transient computation costs approaching 10^4 CPU-hours for a single operating condition [25,28]. Taken together, these ranges span multiple orders of magnitude and define the practical speed–fidelity trade-off: empirical models remain the only viable option for online monitoring, whereas LES is currently restricted to offline mechanism development and validation.

Data needs vs. interpretability

Interpretability determines confidence in engineering decisions, regulatory approval, and diagnostic assessment. As shown in Fig. 3, model transparency decreases as data demands rise, illustrating the trade-off between interpretability and dataset scale.

Empirical and semi-empirical methods (Table 2) use closed-form relationships derived from combustion theory or regression analysis. Their coefficients correspond to identifiable physical quantities, enabling root-cause tracing, validation against measurements, and straightforward auditing. These models remain effective with limited datasets (about 10^2 – 10^3 samples), supporting field calibration,

compliance monitoring, and early-stage design screening.

Purely data-driven approaches, including neural networks, ensemble regressors, and long short-term memory models [37,49], deliver high predictive accuracy when trained on large datasets (roughly 10^4 – 10^5 samples). Their internal parameters lack direct physical meaning, limiting interpretability and diagnostic use. Robustness depends on how well the training data represent the operating domain, and retraining is often required when boundary conditions, fuels, or load regimes vary.

Hybrid frameworks seek to combine accuracy with interpretability by embedding physics-based information into statistical architectures. Examples include models incorporating CFD- or CRN-derived features [2,10], and symbolic-regression formulations that yield algebraic expressions consistent with combustion mechanisms [8]. These maintain partial mechanistic transparency while improving computational efficiency but still depend on careful feature selection and rigorous validation.

This trade-off defines a central research objective: to develop emission models that remain interpretable and data-efficient while providing transferable accuracy across fuels, ambient conditions, and operating transients.

Static vs. adaptive frameworks

Static models are calibrated once and applied without further adjustment. They perform reliably within the tuned operating envelope but lose accuracy as system conditions change. Long-term field studies show degradation from component aging [35], and reduced fidelity under transient or off-design conditions [6,11,29]. These effects intensify in hydrogen-enriched or variable-load combustion, where emissions exhibit strong nonlinearity [23,24].

Adaptive frameworks mitigate these limits through periodic retraining or continuous online learning. Studies using stacked ensembles and hybrid surrogates [2,43,45], sustain accuracy by updating parameters as new operational data are incorporated. Physics-guided adaptive schemes are emerging: Reumschüssel et al. [44] applied Bayesian optimization to refine surrogate fidelity as additional physical data became available, showing how uncertainty metrics can direct targeted data acquisition. These approaches shift practice toward

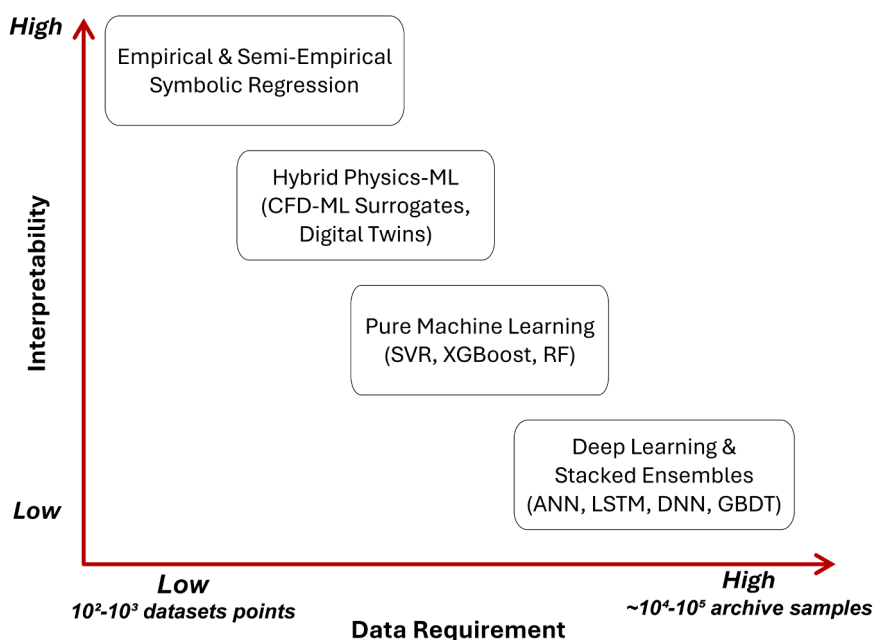


Fig. 3. Data Needs vs. Interpretability.

Conceptual mapping of emission-modeling paradigms showing the trade-off between data requirements and interpretability. Representative dataset scales are drawn from Mubashir et al. [2], Kosheleva and Nikolskii [8], and Dai et al. [49].

models that preserve mechanistic grounding while adapting to evolving conditions.

Adoption of adaptive modeling adds new requirements for data governance and lifecycle management. Continuous learning demands validated datasets, version control, and transparent documentation of retraining cycles to meet regulatory and traceability standards. The shift from static to adaptive frameworks expands predictive capability rather than replacing existing methods. Standardized protocols for drift detection, retraining cadence, and auditability ensure reliability in long-term operation.

For operational deployment, actionable practices can be inferred from validation constraints reported in Lipperheide et al. [35] and Si et al. [37]. Retraining is typically triggered by performance drift, for example when the moving average of the prediction residual exceeds sensor measurement uncertainty for a sustained period. For NO_x, this uncertainty is commonly on the order of $\pm 2\text{--}5$ ppm, consistent with the EPA-compliant precision reported by Si et al. [37], and exceedance over >24 operational hours has been used as a practical threshold. Retraining frequency depends on operating regime: quarterly updates are generally sufficient for base-load units to capture seasonal effects and component aging [35], whereas peaking units or systems operating with hydrogen blends benefit from event-based updating. To maintain stability during online adaptation, parameter updates are commonly constrained within physics-based bounds, a safeguard supported by constrained optimization approaches such as those described by Reumschüssel et al. [44].

Applicability domains and model selection

Model choice reflects the trade-offs summarized in Figs. 2–4 and the application mapping in Table 7. Selection should consider four factors: data availability, computational budget and latency, interpretability requirements, and operating variability.

For rapid design-space exploration, automated CRN pipelines and CFD-derived surrogates are most effective. CRNs enable transparent “what-if” analyses across parametric ranges [4], while polynomial or sparse surrogates built from CFD data can evaluate designs orders of magnitude faster than full simulations while retaining core physics [10]. Use these when coverage and turnaround time outweigh pointwise precision.

For analyzing flame stabilization, heat loss, and radical pathways, high-fidelity physics remains indispensable. A tiered workflow is practical: use LES or finite-rate RANS to anchor boundary conditions and mechanisms [25,28], then apply reduced CRNs or surrogates to map sensitivities. Hybrid LES→CRN or LES→surrogate chains reduce cost while maintaining causal traceability, especially for hydrogen and lean-premixed regimes.

Where auditability is required, interpretable or semi-empirical PEMS are preferred, optionally enhanced with degradation-aware corrections [35]. Black-box models can be acceptable when validation, drift

monitoring, and change-control procedures are documented [37]. Models whose parameters retain physical meaning remain most defensible under regulatory scrutiny.

Millisecond-level inference favors compact ML surrogates or ensembles trained on curated feature sets [41,49]. They should include safeguards such as domain checks, uncertainty thresholds, and fallback logic constrained by physics-based limits. Online retraining is warranted only when data quality and governance are verified.

For fleet benchmarking and life-cycle analytics, physics-informed surrogates coupled with Bayesian or active-learning loops can target data in high-uncertainty regions [44]. Stacked or hybrid surrogates scale efficiently across units but require strict documentation of dataset provenance and split methodology to avoid inflated performance claims [2,45].

In practice, start with the least complex framework that meets the decision objective.

Synthesis of research gaps and future directions

Significant progress has been achieved across empirical, physics-based, machine-learning, and hybrid methods, yet key limitations still restrict predictive reliability under operational conditions.

Transient and off-design emissions

Most models rely on steady data, yet turbines experience rapid transients where thermal and chemical time scales overlap. Ding et al. [29] and Tenango-Pirín et al. [24] observed sharp NO_x deviations during load shifts, while Bajwa et al. [11] and LES studies [23,53] linked these spikes to coherent vortical modes. The unresolved issue is how turbulence intermittency modifies finite-rate chemistry beyond quasi-steady assumptions. Future work should quantify this interaction using time-resolved diagnostics and validated transient datasets.

Fuel-Flexible kinetics and mechanistic generalization

The shift toward hydrogen, ammonia, and syngas reveals the fragility of kinetic schemes tuned for methane. LES and CRN studies [23, 25,34], show that altered radical pools and pressure-dependent branching change NO_x formation routes, reducing cross-fuel predictability. Detailed mechanisms remain computationally expensive, while reduced forms lose accuracy outside their calibration range. The barrier lies in preserving pressure and cross-species coupling during simplification. Future work should develop fuel-agnostic reduced mechanisms validated on unified multi-fuel datasets for consistent modeling across hydrogen-carbon blends and zero-carbon fuels.

Data fidelity and domain coverage

Many ML emission models show high accuracy within limited plant datasets but fail when fuel composition, ambient conditions, or component wear differ from the training domain. Studies by Aslan [42], Hoque et al. [43], and Coelho et al. [48] highlight this gap between internal fit and external validity. The main challenge is domain-aware architecture, not data volume. Reliable frameworks should include out-of-domain detection, physics-based feature constraints, and uncertainty propagation. Open benchmark datasets with controlled variation in load, fuel, and environment [2,44] are needed for reproducible training and consistent cross-domain assessment.

Uncertainty quantification and reliability

Most emission models report deterministic accuracy metrics such as RMSE or R² without conveying confidence bounds, leaving predictive reliability untested. Even recent ML frameworks [49,51] seldom distinguish between uncertainty arising from limited data and that intrinsic to combustion variability. Bayesian surrogates [44], and SVR implementations with confidence intervals [2], prove that formal uncertainty quantification (UQ) can be embedded without excessive cost. The next step is to institutionalize UQ as a standard validation layer,

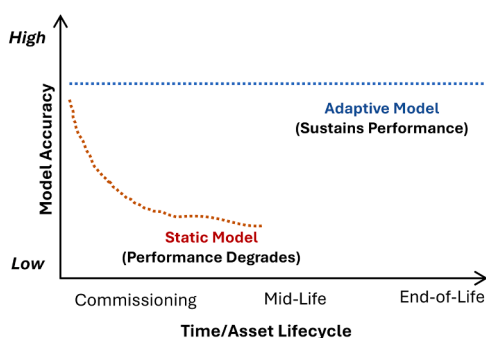


Fig. 4. Static and adaptive emissions modeling across the engine lifecycle. Conceptual schematic illustrating performance drift in static models versus stability in adaptive frameworks. Trends are qualitative and not to scale.

separating epistemic from aleatoric components to ensure traceable, risk-informed deployment of emissions models in operational settings.

Interpretable hybrid architectures

Predictive accuracy alone does not meet diagnostic, certification, or regulatory needs. Hybrid frameworks that combine mechanistic solvers with data-driven inference [2,10] show promise but remain hard to audit. Symbolic regression [8], and physics-informed generative models [56], move toward interpretability, yet scalability across fuels and transients is untested. Effective designs keep latent variables thermochemically meaningful and parameters tied to measurable quantities. Progress relies on shared validation protocols, transparent benchmarks, and open experimental-simulation datasets that support adaptable, fuel-flexible, interpretable models with quantified uncertainty.

Table 6 summarizes the cross-paradigm trade-offs identified in this review, aligning fidelity, interpretability, and adaptability within a unified decision framework. It shows how speed, data scope, and calibration strategy define the operational and research priorities of modern emission-prediction practice, linking the detailed analyses of Sections 3–4 with the conclusions that follow.

Conclusions

No single modeling paradigm provides uniform reliability for gas-turbine emission prediction. Effective practice remains hierarchical, selecting the least complex model that meets the decision requirement and increasing fidelity only when uncertainty, novelty, or regulation demands it, guided by data availability, computational cost, and interpretability.

This review's synthesis provides clear answers to its guiding questions. The surveyed frameworks form a continuum of increasing mechanistic integration, where reliability under transient and fuel-flexible conditions depends primarily on how well a model captures turbulence-chemistry coupling, domain coverage, and explicit uncertainty. The most relevant advancement is the hierarchical convergence of these paradigms, in which physics-based simulations now serve as the mechanistic foundation that trains, constrains, and validates more scalable hybrid frameworks.

Three priorities define the next phase of progress. Validation must emphasize operational realism through transient and off-design benchmarks, with studies reporting hardware-specific runtime, quantified error, domain coverage, and uncertainty. Research should advance fuel-flexible reduced mechanisms validated on systematic multi-fuel datasets to ensure consistent prediction across hydrogen, ammonia, and syngas blends. Future architectures should embed physics-guided features, active learning, and uncertainty quantification so predictions remain

Table 7
Modeling Paradigms and Their High-Impact Use Cases.

Lifecycle Stage	Application Focus	Preferred Modeling Approaches	Primary Advantage
Early Design and Concept Screening	Preliminary combustor geometry, parametric “what-if” studies	Reduced-order CRNs [4]; CFD-derived surrogate models [10]	Enables rapid design exploration with minimal computational cost
Fuel-Flexible and Combustor Development	Hydrogen or SAF combustion, flame stabilization, lean operation	LES with finite-rate chemistry [28]; hybrid LES-CRN coupling [30].	Resolves flame structure and radical pathways for alternative-fuel validation
Regulatory and Compliance Operations	Continuous emissions monitoring and verification	Adaptive hybrid PEMS [35]; validated ANN-based PEMS [37]	Provides audit-compliant accuracy and compensates for operational drift
Real-Time Engine Control and Optimization	Feedback control, transient-load adaptation	Ensemble ML surrogates and LSTM architectures [41, 49]	Achieves millisecond-scale inference for active emissions mitigation
Digital Twin and Fleet-Level Optimization	Multi-engine benchmarking, life-cycle analytics	Integrated physics-ML platforms [44]; automated CRN mapping [4]	Supports predictive maintenance and long-horizon emissions management

both accurate and traceable for engineering and regulatory use.

The evidence synthesized in this review indicates that the next generation of emissions models will not be defined by algorithmic novelty but by their architecture of integration. Future progress will be measured by how effectively these frameworks couple the mechanistic fidelity of physics-based simulations with the adaptive efficiency of data-driven methods. The emerging paradigm is one where high-fidelity physics is no longer the endpoint of modeling but the foundational source of constraints and validation data for more tractable, interpretable, and continuously updated hybrid systems. This shift reframes the ultimate goal: from finding the single 'best' model to engineering a verifiable, lifecycle-aware modeling ecosystem.

CRedit authorship contribution statement

Tariq R. Aboalhol: Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Khaled A. Abdusamad: Visualization, Investigation.

Table 6
Key Trade-Offs Across Modeling Paradigms.

Trade-Off	Low-Fidelity / High-Speed	Physics-ML Hybrids (Intermediate)	High-Fidelity / Low-Speed	Implications (Practice + Research)
Speed vs Fidelity	Semi-empirical correlations and reduced CRNs; run in milliseconds to minutes; simplified transport and kinetics [6,11,12].	Surrogates trained on CFD or rig data; seconds to minutes; retain key physical behavior [2,10,44].	LES or detailed-chemistry CFD; hours to thousands of CPU hours; fully resolve turbulence-chemistry [25,28,52,53].	Practice: Low- and mid-fidelity models enable real-time monitoring and design scans; high-fidelity remains essential for mechanism analysis and novel fuels. Research: Quantify surrogate error bounds and define substitution limits for CFD.
Data Need vs Interpretability	Symbolic or empirical models; transparent; modest data; accuracy declines beyond calibration [6,8].	Physics-guided ML and CFD-trained surrogates; moderate data; partially interpretable through embedded physics [10, 44]	First-principles CFD; inherently interpretable; minimal prior data but high computational cost [28, 52].	Practice: Transparency aids auditing and regulatory traceability. Research: Develop explainable and uncertainty-aware hybrid frameworks benchmarked against interpretable baselines.
Static Calibration vs Adaptive Updating	Fixed correlations or lookup tables; performance drifts with aging, fuel shifts, or ambient change [11].	Periodic or online retraining using operational data; accuracy maintained if data pipelines remain reliable [41,45].	High-fidelity models generalize well but are too slow for frequent updates [25,28].	Practice: Fleet monitoring benefits from hybrids with scheduled retraining and safeguards. Research: Formalize data-curation, drift-detection, and safe-update protocols for dynamic operation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Q. Huang, R. Yang, J. Liu, T. Xie, J. Liu, Investigation of the mechanism behind the surge in nitrogen dioxide emissions in engines transitioning from pure diesel operation to methanol/diesel dual-fuel operation, *Fuel Process. Technol.* 264 (2024) 108131, <https://doi.org/10.1016/j.fuproc.2024.108131>.
- [2] M. Mubashir, D. Shen, H. Kraiem, A. Flah, N.F. Alshammari, M.M. Hanif, Machine learning assisted CFD optimization of fuel-staging natural gas burners for enhanced combustion efficiency and reduced NOx emissions, *Sci. Rep.* 15 (1) (2025) 23547, <https://doi.org/10.1038/s41598-025-05132-8>.
- [3] R. Ferrari, Writing narrative style literature reviews, *Med. Writ.* 24 (4) (2015) 230–235, <https://doi.org/10.1179/2047480615Z.000000000329>.
- [4] V. Fichet, M. Kanniche, P. Plion, O. Gicquel, A reactor network model for predicting NOx emissions in gas turbines, *Fuel* 89 (9) (2010) 2202–2210, <https://doi.org/10.1016/j.fuel.2010.02.010>.
- [5] N. Chandrasekaran, A. Guha, Study of prediction methods for NOx emission from turbofan engines, *J. Propul. Power* 28 (1) (2012) 170–180, <https://doi.org/10.2514/1.B34245>.
- [6] Z. Xi, Z. Fu, X. Hu, S.W. Sabir, Y. Jiang, An experimental investigation on the NO and CO emission characteristics of a swirl convergent-divergent nozzle at elevated pressure, *Energies* (Basel) 11 (6) (2018) 1410, <https://doi.org/10.3390/en11061410>.
- [7] W. Katzenstein, J. Apt, Air emissions due to wind and solar power, *Environ. Sci. Technol.* 43 (2) (2009) 253–258, <https://doi.org/10.1021/es801437t>.
- [8] O. Kochueva, K. Nikolskii, Data analysis and symbolic regression models for predicting CO and NOx emissions from gas turbines, *Computation* 9 (12) (2021) 139, <https://doi.org/10.3390/computation9120139>.
- [9] Q. Zhao, G. Wang, H. Zhang, Y. Xu, S. Yang, Estimation of NOx emissions from the combustion chamber of heavy-duty gas turbines, *J. Phys. Conf. Ser.* 657 (1) (2021) 012004, <https://doi.org/10.1088/1755-1315/657/1/012004>.
- [10] Q. Zhao, F. Liu, A. Jiao, Q. Yang, H. Xu, X. Liao, Prediction model of NOx emissions in the heavy-duty gas turbine combustor based on MILD combustion, *Energy* 282 (2023) 128974, <https://doi.org/10.1016/j.energy.2023.128974>.
- [11] A. Bajwa, G. Zou, F. Zhong, X. Fang, F. Leach, M. Davy, Development of a semi-empirical physical model for transient NOx emissions prediction from a high-speed diesel engine, *Int. J. Engine Res.* 25 (10) (2024) 1835–1848, <https://doi.org/10.1177/14680874241255165>.
- [12] H.O. Egbare, C.C. Kwasi-Effah, A novel empirical model for predicting the carbon dioxide emission of a gas turbine power plant, *Heliyon* 9 (3) (2023) e14645, <https://doi.org/10.1016/j.heliyon.2023.e14645>.
- [13] Q. Huang, T. Xie, J. Liu, Machine learning-assisted reconstruction of In-cylinder pressure in internal combustion engines under unmeasured operating conditions, *Energies* 18 (19) (2025) 5235, <https://doi.org/10.3390/en18195235>.
- [14] I.V. Novoselov, P.C. Malte, Development and application of an eight-step global mechanism for CFD and CRN simulations of lean-premixed combustors, in: *Proceeding ASME Turbo Expo 47918*, 2007, pp. 769–779, <https://doi.org/10.1115/GT2007-27990>.
- [15] J. Park, T.H. Nguyen, D. Joung, K.Y. Huh, M.C. Lee, Prediction of NO x and CO emissions from an industrial lean-premixed gas turbine combustor using a chemical reactor network model, *Energy Fuels* 27 (3) (2013) 1643–1651, <https://doi.org/10.1021/ef301741t>.
- [16] P.S. Filippov, P.Y. Khudyakov, A.F. Ryzhkov, Validation of the thermal NOx emissions model from a gas fuel combustor under atmospheric pressure, *J. Phys.: Conf. Ser.* 899 (9) (2017) 092005, <https://doi.org/10.1088/1742-6596/899/9/092005>.
- [17] A. Innocenti, A. Andreini, D. Bertini, B. Facchini, M. Motta, Turbulent flow-field effects in a hybrid CFD-CRN model for the prediction of NOx and CO emissions in aero-engine combustors, *Fuel* 215 (2018) 853–864, <https://doi.org/10.1016/j.fuel.2017.11.097>.
- [18] J.A. Van Oijen, L.P.H. De Goeij, Modelling of premixed counterflow flames using the flamelet-generated manifold method, *Combust. Theory Model.* 6 (3) (2002) 463, <https://doi.org/10.1088/1364-7830/6/3/305>.
- [19] G. Goldin, Z. Ren, H. Forkel, L. Lu, V. Tangirala, H. Karim, Modeling CO with flamelet-generated manifolds: part 1-flamelet configuration, *Proceeding ASME Turbo Expo*, Am. Soc. Mech. Eng. 44687 (2012) 1141–1151, <https://doi.org/10.1115/GT2012-69528>.
- [20] N. Klarmann, T. Sattelmayer, W. Geng, F. Magni, Flamelet generated manifolds for partially premixed, highly stretched and non-adiabatic combustion in gas turbines, in: *54th AIAA Aerospace Sciences Meeting* (p. 2120), 2016, <https://doi.org/10.2514/6.2016-2120>.
- [21] N. Klarmann, B.T. Zoller, T. Sattelmayer, Numerical modeling of CO-emissions for gas turbine combustors operating at part-load conditions, *J. Glob. Power Propuls. Soc.* 2 (2018) 376–387, <https://doi.org/10.22261/C3N50A>.
- [22] A.A. Perpignan, M.G. Talboom, Y. Levy, A.G. Rao, Emission modeling of an interturbine burner based on flameless combustion, *Energy Fuels* 32 (1) (2018) 822–838, <https://doi.org/10.1021/acs.energyfuels.7b02473>.
- [23] Y. Wang, J. Liu, L. Wang, Z. Fu, P. Weng, Non-premixed combustion and NOx emission characteristics in a micro gas turbine swirl combustor fueled by methane and ammonia at various heat loads, *Heliyon* 9 (3) (2023), <https://doi.org/10.1016/j.heliyon.2023.e14521>.
- [24] O. Tenango-Pirfin, S. Espinosa, J.C. García, J.A. Rodríguez, Analysis of gas turbine combustor exhaust emissions: effects of transient inlet air pressure, *Energy Technol.* 12 (3) (2023) 2301093, <https://doi.org/10.1002/ente.202301093>.
- [25] S. Castellani, P.C. Nassini, A. Andreini, R. Meloni, E. Pucci, A. Valera-Medina, et al., Numerical modelling of swirl stabilised lean-premixed H₂-CH₄ flames with the artificially thickened flame model, in: *Proceeding ASME Turbo Expo 86953*, 2023, <https://doi.org/10.1115/GT2023-101994>. V03AT04A044.
- [26] E. Fooladgar, C. Brackmann, M. Mannazhi, Y. Ögren, P.E. Bengtsson, H. Wiinikka, P. Tóth, CFD modeling of pyrolysis oil combustion using finite rate chemistry, *Fuel* 299 (2021) 120856, <https://doi.org/10.1016/j.fuel.2021.120856>.
- [27] M. Hajivand, High-fidelity RANS CFD simulations of physico-chemical process of combustion in gas turbine combustion chambers in ANSYS CFX, *Energy Eng. Control Syst.* 10 (2) (2024) 81–95, <https://doi.org/10.23939/jeecs2024.02.081>.
- [28] A. Ercole, D. Lörstad, C. Fureby, Large Eddy simulations of a turbulent premixed swirling flame with finite-rate chemistry and flame-wrinkling turbulent combustion models, *Flow. Turbul. Combust.* (2025) 1–28, <https://doi.org/10.1007/s10494-025-00652-9>.
- [29] Y. Ding, J. Hao, A. Li, X. Wang, X. Zhang, Y. Liu, Numerical simulation of combustion and emission characteristics during gas turbine startup procedure, *Energies* 15 (15) (2022) 5444, <https://doi.org/10.3390/en15155444>.
- [30] A.A. Beige, A. Mardani, An investigation on flame structure and NOx formation in a gas turbine model combustor using large eddy simulation, *Phys. Fluids* 35 (7) (2023) 075146, <https://doi.org/10.1063/5.0155974>.
- [31] A.M. Starik, V.E. Kozlov, A.B. Lebedev, N.S. Titova, Application of reactor net models for the simulation of gas-turbine combustor emissions, *Int. J. Sustain. Aviat.* 1 (1) (2014) 43–57, <https://doi.org/10.1504/IJSA.2014.062867>.
- [32] B. Wegner, U. Gruschka, W. Krebs, Y. Egorov, H. Forkel, J. Ferreira, K. Aschmoneit, CFD prediction of portload CO emissions using a two-timescale combustion model, *J. Eng. Gas Turbines Power* 133 (7) (2011) 071502, <https://doi.org/10.1115/1.4400201>.
- [33] A. Andreini, C. Bianchini, G. Cacioli, B. Facchini, A. Giusti, F. Turrini, Multi-coupled numerical analysis of advanced lean burn injection systems, in: *Proceeding ASME Turbo Expo 45691*, 2014, <https://doi.org/10.1115/GT2014-26808>. V04BT04A041.
- [34] L. Cheng, Y. Chen, Y. Pei, G. Sun, J. Zou, S. Peng, Y. Zhang, NO and CO emission characteristics of laminar and turbulent counterflow premixed hydrogen-rich syngas/air flames, *Processes* 12 (3) (2024) 475, <https://doi.org/10.3390/pr12030475>.
- [35] M. Lipperheide, F. Weidner, M. Wirsum, M. Gassner, S. Bernero, Long-term NOx emission behavior of heavy duty gas turbines: an approach for model-based monitoring and diagnostics, in: *Proceeding ASME Turbo Expo 50916*, 2017, <https://doi.org/10.1115/GT2017-63181>. V006T05A001.
- [36] H. Kaya, P. Tüfekci, E. Uzun, Predicting CO and NOx emissions from gas turbines: novel data and a benchmark PEMS, *Turk. J. Electr. Comput. Sci.* 27 (6) (2019) 4783–4796, <https://doi.org/10.3906/elk-1807-87>.
- [37] M. Si, T.J. Tarnocz, B.M. Wiens, K. Du, Development of predictive emissions monitoring system using open source machine learning library-keras: a case study on a cogeneration unit, *IEE Access* 7 (2019) 113463–113475, <https://doi.org/10.1109/ACCESS.2019.2930555>.
- [38] I. Niño-Adán, E. Portillo, I. Landa-Torres, D. Manjarres, Normalization influence on ANN-based models performance: a new proposal for features' contribution analysis, *IEE Access* 9 (2021) 125462–125477, <https://doi.org/10.1109/ACCESS.2021.3110647>.
- [39] M. Dirik, Prediction of NOx emissions from gas turbines of a combined cycle power plant using an ANFIS model optimized by GA, *Fuel* 321 (2022) 124037, <https://doi.org/10.1016/j.fuel.2022.124037>.
- [40] S. Li, H. Zhu, M. Zhu, G. Zhao, X. Wei, Combustion tuning for a gas turbine power plant using data-driven and machine learning approach, *J. Eng. Gas Turbines Power* 143 (3) (2021) 031021, <https://doi.org/10.1115/GT2020-15788>.
- [41] R. Potts, R. Hackney, G. Leontidis, Tabular machine learning methods for predicting gas turbine emissions, *Mach. Learn. Knowl. Extr.* 5 (3) (2023) 1055–1075, <https://doi.org/10.3390/make5030055>.
- [42] E. Aslan, Prediction and comparative analysis of emissions from gas turbines using random search optimization and different machine learning-based algorithms, *Bull. Pol. Acad. Sci. Tech. Sci.* 72 (6) (2024) e151956, <https://doi.org/10.24425/bpasts.2024.151956>.
- [43] K.E. Hoque, T. Hossain, A.M. Haque, M.A.K. Miah, M.A. Haque, NOx emission predictions in gas turbines through integrated data-driven machine learning approaches, *J. Energy Resour. Technol.* 146 (7) (2024) 071201, <https://doi.org/10.1115/1.4065200>.
- [44] J.M. Reumschüssel, J.G. von Saldern, B. Čosić, C.O. Paschereit, Data-driven optimization of a gas turbine combustor: a Bayesian approach addressing NOx emissions, lean extinction limits, and thermoacoustic stability, *Data-Centric Eng.* 5 (2024) e32, <https://doi.org/10.1017/dce.2024.29>.
- [45] R. Winarto, M.H. Purnomo, W. Anggraeni, Enhancing the performance of gas turbines predictive emissions monitoring system: a stacked ensemble machine learning models, in: *Proceeding IEEE CIVEMSA*, 2025, pp. 1–6, <https://doi.org/10.1109/CIVEMSA65862.2025.11084825>.
- [46] R. Aquize, A. Cajahuaringa, J. Machuca, D. Mauricio, J.M. Mauricio Villanueva, System identification methodology of a gas turbine based on artificial recurrent neural networks, *Sensors* 23 (4) (2023) 2231, <https://doi.org/10.3390/s23042231>.

- [47] M. Faqih, M.B. Omar, R. Ibrahim, Prediction of dry-low emission gas turbine operating range from emission concentration using semi-supervised learning, *Sensors* 23 (8) (2023) 3863, <https://doi.org/10.3390/s23083863>.
- [48] L. dos Santos Coelho, H.V.H. Ayala, V.C. Mariani, CO and NOx emissions prediction in gas turbine using a novel modeling pipeline based on the combination of deep forest regressor and feature engineering, *Fuel* 355 (2024) 129366, <https://doi.org/10.1016/j.fuel.2023.129366>.
- [49] H. Dai, T. Zhao, J. Cao, P. Li, Soft sensor modeling method and application based on TSECIT2FNN-LSTM, *Sci. Rep.* 14 (1) (2024) 23709, <https://doi.org/10.1038/s41598-024-75009-9>.
- [50] N.B. Shaik, K. Jongkittinarukorn, K. Bingi, XGBoost based enhanced predictive model for handling missing input parameters: a case study on gas turbine. *Case stud. Chem. Environ. Eng.* 10 (2024) 100775, <https://doi.org/10.1016/j.csee.2024.100775>.
- [51] V. Sinap, Outliers treatment for improved prediction of CO and NOx emissions from gas turbines using ensemble regressor approaches, *J. Intell. Syst. Theory Appl.* 8 (1) (2025) 63–83, <https://doi.org/10.38016/jista.1566965>.
- [52] K. Yunoki, T. Nishiie, R. Kurose, Prediction of CO emissions in turbulent super lean premixed combustion under pressurized conditions using an LES/non-adiabatic FGM approach, *Int. J. Gas Turbine Propul. Power Syst.* 12 (3) (2021) 1–8, <https://doi.org/10.38036/jgpp.12.3.1>.
- [53] L.M. Chikishev, D.K. Sharaborin, A.S. Lobasov, A.A. Dekterev, R.V. Tolstoguzov, V. M. Dulin, D.M. Markovich, LES simulation of a model gas-turbine lean combustor: impact of coherent flow structures on the temperature field and concentration of CO and NO, *Energies*. (Basel) 15 (12) (2022) 4362, <https://doi.org/10.3390/en15124362>.
- [54] F. Biagioli, F. Güthe, Effect of pressure and fuel-air unmixedness on NOx emissions from industrial gas turbine burners, *Combust. Flame* 151 (1–2) (2007) 274–288, <https://doi.org/10.1016/j.combustflame.2007.04.007>.
- [55] S. Serbin, B. Diasamidze, M. Dzida, Investigations of the working process in a dual-fuel low-emission combustion chamber for an FPSO gas turbine engine, *Pol. Marit. Res.* (3) (2020) 89–99, <https://doi.org/10.2478/pomr-2020-0050>.
- [56] M. Bode, M. Gauding, D. Goeb, T. Falkenstein, H. Pitsch, Applying physics-informed enhanced super-resolution generative adversarial networks to turbulent premixed combustion and engine-like flame kernel direct numerical simulation data, *Proc. Combust. Inst.* 39 (4) (2023) 5289–5298, <https://doi.org/10.1016/j.proci.2022.07.254>.
- [57] L. Zhou, W. Wang, LES of propane-air swirling non-premixed flame using a SOM combustion model, *Combust. Engines* 200 (1) (2025) 71–77, <https://doi.org/10.19206/CE-197157>.