Machine Learning for CH_4/H_2 combustion: Developing high-precision and interpretable surrogate models

Solmaz Nadiri[1]*, Ravi Fernandes[1]
*lead presenter: Solmaz.nadiri@ptb.de
1 Physikalisch-Technische Bundesanstalt (PTB), Germany

The transition to sustainable energy systems demands cleaner and more efficient combustion technologies, with CH_4/H_2 blends offering a promising low-carbon alternative. However, accurately predicting their combustion performance and emissions across diverse operating conditions remains challenging due to the complex, nonlinear nature of combustion reactions. This study employs machine learning (ML) to develop high-fidelity surrogate models for CH_4/H_2 combustion, enabling accurate predictions of key performance indicators such as ignition delay time (IDT), combustion efficiency, and emissions.

A comprehensive dataset was generated using constant volume reactor simulations implemented in Cantera, incorporating the NUIGMech1.3 [1] detailed chemical kinetic mechanism. The dataset covers a broad range of conditions: initial temperatures from 800 K to 1600 K, pressures of 1, 10, 20, and 40 bar, equivalence ratios of 0.5, 1.0, and 2.0, and fuel compositions with H_2 content varying from 0% to 100% in increments of 20%. The role of oxidizer humidity was also examined at 0, 0.05, and 0.1 mole fractions to assess its effect on combustion properties. The simulations resulted in 1,080 unique cases, capturing critical combustion characteristics under diverse conditions.

The output variables encompass IDT, combustion efficiency, and pollutant emissions including CO_2, CO, NO, NO_2, N_2O, and CH_2O. Several ML models were trained and evaluated, including CatBoost, XGBoost, Random Forest, Gaussian Process Regression (GPR), and Neural Networks. The results demonstrate that gradient boosting models, particularly CatBoost, achieve the highest predictive accuracy, with R^2 values exceeding 0.99 for most outputs. GPR provided valuable uncertainty quantification, ensuring model reliability in scenarios where confidence intervals are required.

To enhance model interpretability, SHapley Additive exPlanations (SHAP) analysis was applied, revealing that temperature, equivalence ratio, and H_2 content are the dominant factors affecting combustion behavior. The findings confirm that higher H2 content leads to lower CO and CO_2 emissions but increases NO_x formation due to elevated flame temperatures. Furthermore, the developed surrogate models were used in an optimization framework employing Genetic Algorithms (GA) to identify optimal operating conditions that balance efficiency and emissions. The optimal equivalence ratio (~0.67) and moderate H_2 enrichment (~33%) resulted in near-complete combustion (99.9% efficiency) while maintaining low NO_x emissions (<0.0056) and minimal CO formation.

References

[1] S. Dong. et al. On the low-temperature chemistry of 1, 3-butadiene. Proceedings of the Combustion Institute 39, 365–373 (2023).