FEDERAL UNIVERSITY OF SANTA CATARINA
TECHNOLOGICAL CENTRE OF JOINVILLE
AEROSPACE ENGINEERING

BEATRIZ LETÍCIA BINOTTO

CLUSTERING DATA VIA REACTION NETWORK COMPARISON: A PRIORI IDENTIFICATION OF SIMULATION TARGETS FOR GRAPH-BASED REDUCTION TECHNIQUES

Joinville
2024

BEATRIZ LETÍCIA BINOTTO

CLUSTERING DATA VIA REACTION NETWORK COMPARISON: A PRIORI IDENTIFICATION OF SIMULATION TARGETS FOR GRAPH-BASED REDUCTION TECHNIQUES

Thesis presented as a requirement for obtaining a Bachelor's degree in Aerospace Engineering at the Technological Centre of Joinville of the Federal University of Santa Catarina.

First supervisor: Prof. Dr. Eng. Rafael Catapan

Second supervisor: Prof. Dr.-Ing. Heinz Pitsch

Joinville
2024

BEATRIZ LETÍCIA BINOTTO

CLUSTERING DATA VIA REACTION NETWORK COMPARISON: A PRIORI IDENTIFICATION OF SIMULATION TARGETS FOR GRAPH-BASED REDUCTION TECHNIQUES

This thesis was considered eligible for obtaining a Bachelor's degree in Aerospace Engineering at the Technological Centre of Joinville of the Federal University of Santa Catarina.

Joinville (SC), 06 December 2024.

Examining Board:

Prof. Dr. Eng. Rafael Catapan
First Supervisor/Board Chairman
UFSC

Francesca Loffredo

Firmato digitalmente da Francesca Loffredo
Data: 2024.12.18 10:57:32 +01'00'

Francesca Loffredo M.Sc.
ITV Supervisor
ITV RWTH

Prof. Dr. Eng. Leonel R Cancino
Member
UFSC

CONFIDENTIALITY STATEMENT

This bachelor thesis, titled *Clustering Data via Reaction Network Comparison: A Priori Identification of Simulation Targets for Graph-Based Reduction Techniques*, was completed during my internship at the Institute for Combustion Technology (ITV) at RWTH Aachen University.

The work presented herein, including all results, methodologies, and knowledge derived, is the exclusive property of ITV.

Disclosure, reproduction, or use of this content is prohibited without prior written permission from ITV. A violation of this agreement can result in legal liability.

Documento assinado digitalmente
Beatriz Leticia Binotto
Data: 18/12/2024 07:50:53-0300
CPF: ***.394.039-**
Verifique as assinaturas em https://v.ufsc.br

Beatriz Letícia Binotto
UFSC

Francesca Loffredo

Firmato digitalmente
da Francesca Loffredo
Data: 2024.12.18
10:58:06 +01'00'

Francesca Loffredo M.Sc.
ITV RWTH

To Marcello Carnauba Carpeggiani, who started this journey with me, but did not live to see it finished.

ACKNOWLEDGEMENTS

To my parents, for always supporting my studies, and to my sister, for being my inspiration.

To Patrick, for the partnership and constant encouragement.

To my supervisor, Professor Catapan, and Professor Cancino, for the guidance and valuable contributions.

To Professor Pitsch and Francesca, for the opportunity to develop this work at ITV and support.

To Lígia, Gabriel, and João, for sharing the university journey with me.

To André, Júlia, and Lukas, for making Germany feel as warm as Brazil.

To Thaís, Matheus, Enrico, Pedro, Alexandre, and Gustavo, for being my constant connection to home.

“ Somewhere, something incredible is waiting to be known. ”

Sharon Begley

ABSTRACT

The use of detailed chemical models in combustion simulations allows for an accurate representation of reactive flows. Because of the growing complexity of kinetic models, reduction strategies aimed at developing simplified kinetic schemes, while maintaining high predictive accuracy, have become increasingly important for enabling large-scale simulations that would otherwise be computationally costly. Graph-based reduction approaches rely on the definition of physically meaningful coefficients that quantify the coupling strength between species. These coefficients are generally determined through the analysis of a database of numerically computed solutions, chosen as sampled representative conditions of the applicability domain of the detailed kinetic model. However, this sampling is often based on expert knowledge, and its effects on the reduction process are frequently neglected. In this work, a systematic approach for selecting simulation targets for graph-based reduction techniques is defined. Starting with a database of computed solutions representative of 264 ignition delay time experiments for ammonia (NH_3) combustion, an affinity propagation clustering algorithm is used to group experiments based on the similarity of their reaction networks. The experiments are grouped into 17 clusters and a representative center is chosen from each cluster to serve as a simulation target during the reduction phase. The performances of the obtained reduced models are evaluated through quantitative assessment against results obtained with the detailed model. Further, the effectiveness and limitations of the proposed method are discussed. First, the capability of the clustering algorithm in detecting variation in the underlying chemistry representative of different combustion regimes is illustrated. Secondly, a comparison between reduced models obtained by ranking the species based on the coefficients of importance obtained by considering only the centers against those obtained with all the cases is presented.

Keywords: Chemical kinetic models. Graph-Based Reduction. Ammonia-hydrogen combustion.

RESUMO

O uso de modelos químicos detalhados em simulações de combustão possibilita uma representação precisa dos fluxos reativos. Com o aumento da complexidade dos modelos cinéticos, as estratégias de redução, que buscam simplificar os esquemas cinéticos sem comprometer a precisão, se tornaram cada vez mais importantes. Essas estratégias tornam viáveis simulações em grande escala, que, de outro modo, teriam um custo computacional muito elevado. As abordagens de redução baseadas em gráficos dependem da definição de coeficientes que quantificam a força de acoplamento entre as espécies. Em geral, esses coeficientes são obtidos por meio de análise de um banco de dados contendo soluções computadas numericamente, selecionadas para refletir condições típicas de aplicação do modelo cinético detalhado. No entanto, essa seleção de amostras costuma exigir um conhecimento especializado, e os impactos dessa escolha no processo de redução nem sempre são avaliados. Este trabalho apresenta uma abordagem sistemática para a seleção de alvos de simulação em técnicas de redução baseadas em gráficos. A partir de um banco de dados contendo soluções que representam 264 experimentos de tempo de atraso de ignição para a combustão de amônia (NH_3), é utilizado um algoritmo de clustering por afinidade de propagação para agrupar os experimentos com base na similaridade de suas redes de reação. Os experimentos são agrupados em 17 clusters, e o centro de cada cluster é selecionado como alvo de simulação durante a fase de redução. O desempenho dos modelos reduzidos é avaliado de forma quantitativa, por meio da comparação com os resultados obtidos pelo modelo detalhado. Adicionalmente, a eficácia e as limitações do método proposto são discutidas. Primeiro, é demonstrada a capacidade do algoritmo de clustering em identificar variação na química básica representativa de diferentes regimes de combustão. Em seguida, é realizada uma comparação entre os modelos reduzidos obtidos pelo ranking das espécies com base nos coeficientes de importância considerando apenas os centros, e os resultados obtidos ao incluir todos os casos.

Palavras-chave: Modelos Cinéticos. Redução baseada em gráficos. Combustão de amônia-hidrogênio.

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

Symbol	Definition
CM	Curve Matching
CSP	Computational Singular Perturbation
DRG	Directed Relation Graph
DRGEP	Directed Relation Graph with Error Propagation
DRGSR	Directed Relation Graph-based SpeciesRank
IDT	Ignition Delay time
JSD	Jensen-Shannon Divergence
KLD	Kullback-Leibler Divergence
NASA	National Aeronautics and Space Administration
PC	Principal Component
PCA	Principal Component Analysis
PFA	Path Flux Analysis
QSSA	Quasi-Steady-State Assumptions
RCM	Rapid Compression Machine
SSE	Sum Squared Error
ST	Shock Tube

LIST OF SYMBOLS

Symbol	Definition
$\mathbf{A}$	Adjacency matrix
A	Frequency factor
a_i	NASA ith numerical coefficient
C	total concentration
C_k	Consumption of species k
c_P	Specific heat at constant pressure
c_V	Specific heat at constant volume
$d^0_{L^2}$	CM 1st dissimilarity
$d^1_{L^2}$	CM 2nd dissimilarity
d^0_{Pe}	CM 3rd dissimilarity
d^1_{Pe}	CM 4th dissimilarity
E_A	Activation energy
G	Gibbs free energy
h	Enthalpy
K_C	Chemical equilibrium constant
k_b	Backward reaction rate coefficient
k_f	Forward reaction rate coefficient
$\dot{m}$	Chemical production rate
M	Molecular Weight
$\mathsf{n_E}$	Number of educts
$\mathsf{n_{Pr}}$	Number of products
n_R	Number of reversible reactions
$\mathbf{p}$	Preference vector
P	Pressure
P_k	Production of species k
$\mathbf{pr}$	Ranking vector
Q	Heat
r_{AB}	Path-dependent interaction coefficient
R_u	Universal gas constant
$\mathbf{S}$	Similarity matrix
T	Temperature
t	Time
V	Volume

continued on next page

Symbol	Definition
v'	Stoichiometric coefficients of reactants
v''	Stoichiometric coefficients of products
Y	Mass fraction
ρ	Density
Φ	Equivalence ratio
ω	Reaction Rate

CONTENTS

1 INTRODUCTION

Since the Industrial Revolution, human activity has emitted a significant amount of carbon dioxide and greenhouse gases into the atmosphere, resulting in serious climate change (LEE et al., 2022). To overcome this environmental crisis, many countries have begun to focus on strategies to reduce carbon dioxide emissions and expand the supply of renewable energy.

Fossil fuels are responsible for more than 75 percent of global greenhouse gas emissions and almost 90 percent of all carbon dioxide emissions (United Nations, 2024). The primary source of greenhouse gas emissions is the combustion of fossil fuels for electricity, heat, and transport (RITCHIE et al., 2020). In 2019, direct aviation emissions accounted for 2.5% of global CO_2 emissions, and have contributed 4% to the global temperature rise since pre-industrial times (RITCHIE, 2024).

Clean ammonia is considered a potential solution that can address the challenges associated with hydrogen energy. With a molecular formula of NH_3, ammonia has a hydrogen content of 17.6 wt% and contains no carbon. It enables the storage, transportation, and distribution of hydrogen without requiring significant changes to the existing infrastructure currently deployed by industries (ERDEMIR; DINCER, 2021).

Being one of the most widely produced industrial chemicals globally (GIDDEY et al., 2013), the use of ammonia can play a role in achieving a faster transition to a hydrogen economy through its use in combustion processes, such as internal combustion engines and gas turbines, with minimal modifications needed in existing industrial infrastructure (ERDEMIR; DINCER, 2021).

Ammonia is not a new concept in the aerospace industry. In the 1960s, the X-15 aircraft reached a speed of Mach 6.7, setting a record for crewed, powered aircraft. Its rocket engine used ammonia and liquid oxygen as propellants (NASA, 2024). Today, ammonia is regaining attention as a candidate for sustainable aviation fuel.

In a retrofit study of an Airbus A320neo, modifications were made to the propulsion subsystem to replace traditional kerosene with ammonia (RAO et al., 2022). The integration of a cracking system, which uses a catalyst to break some of the ammonia into hydrogen, allows hydrogen to be mixed with the remaining ammonia to create a fuel blend with improved combustion characteristics. This ammonia-hydrogen mixture is fed into the engine and burns like conventional fuel, enabling existing engines and aircraft designs to be adapted for its use (BLAIN, 2022).

Developing experiments to study combustion in engines and burners is challenging due to high implementation costs and development complexity (GARCIA; SANTOS, 2022). As a result, the application of detailed reaction mechanisms in combustion simulations has been extensively used to provide accurate descriptions of chemically

reactive flows. Although the available computational power is growing, the implementation of detailed mechanisms in numerical simulations requires significant computation time due to the combustion process involving hundreds of species and thousands of reactions.

To address these computational demands, two primary approaches exist: time-scale analysis, such as quasi-steady-state assumptions (QSSA) or computational singular perturbation (CSP), and graph-based reduction (PEPIOT-DESJARDINS; PITSCH, 2008). The latter approaches simplification by reducing the number of species in the mechanism without compromising the accuracy of key properties.

Graph-based methods are built upon the schematization of the reaction network in the form of a directed-weighted graph, being an appealing method due to the low computational costs involved in their applications compared to other techniques. Once the species and reactions with minimal contributions to a particular application are removed through graph-based skeletal reduction, the resulting mechanism can be further simplified through the time-scale analysis.

Directed Relation Graph (DRG) (LU; LAW, 2006), DRG with Error Propagation (DRGEP) (PEPIOT-DESJARDINS; PITSCH, 2008), and Path Flux Analysis (PFA) (SUN et al., 2010) are widely used techniques for skeletal mechanism reduction. All three methods define coefficients to estimate the error in the overall production or consumption of a species if another species is excluded from the model, where a cutoff threshold is then applied to the system to eliminate species and reactions. The approaches differ in their definition of the coefficients and how the indirect relationships between species are treated.

In all these models, the coefficients are computed using a database of numerically computed solutions appropriately sampled to encompass the validity range of the detailed model. Constructing such a database involves sampling the thermodynamic space, defined by pressure, temperature, and equivalence ratio, within the validity range of the detailed model. The reduced model is then assumed to be valid in the vicinity of these sampled cases (PEPIOT-DESJARDINS; PITSCH, 2008). However, the sampling is often performed based on expert knowledge, with little consideration given to its impact on the reduction process. To the author's knowledge, a method to systematically determine the validation targets has not been proposed yet.

In this thesis, the effect of defining appropriate validation targets for graph-based reduction techniques is evaluated. Based on the assumption that similar reaction networks will share the same simplified scheme, a method aiming at identifying similarities between different thermodynamic states is developed. The validity of the proposed approach is verified on a wide range of initial conditions, which are derived from 264 shock tube experiments.

1.1 OBJECTIVES

1.1.1 General Objective

To develop a method for the automated identification of simulation targets for chemical kinetic mechanism reduction through a quantitative comparison of reaction networks.

1.1.2 Specific Objectives

- To review foundational principles in chemical kinetics.
- To conduct a literature review on current reduction techniques and their challenges.
- To assess the validity and applicability of the proposed algorithm.

2 THEORETICAL BACKGROUND

In this section, the theoretical background supporting this work is provided.

The discussion begins with an overview of ammonia blends, as the database used in this study focuses on ammonia and ammonia-hydrogen mixtures. Subsequently, the concept of ignition delay time is introduced. Experimental facilities employed to measure this quantity are introduced, as well as the numerical models used to simulate the experimental apparatus. The governing equations that support the simulations are addressed in subsection 2.4.1.

Chemical kinetic combustion models are discussed next, in section 2.5, covering the differentiation between global and elementary reactions, and the determination of reaction rates.

Lastly, existing graph-based reduction techniques are reviewed in section 2.6, emphasizing their role in simplifying complex chemical kinetic mechanisms without significant loss in predicting relevant combustion properties.

2.1 AMMONIA BLENDS

The use of ammonia as a fuel has several advantages (ERDEMIR; DINCER, 2021), summarized below:

- It is a carbon-free fuel.
- It is a potential hydrogen storage and carrier.
- It can be produced from renewable sources.
- Compared to hydrogen, ammonia transportation is safer and more practical.
- Its distinctive odor facilitates easy detection of leaks.
- For its use in engines and gas turbines, only minor modifications are required.

Although ammonia exhibits low reactivity in its pure form, with high ignition temperature (ERDEMIR; DINCER, 2021), this issue can be overcome by blending ammonia with traditional fuels such as gasoline, diesel, ethanol, methanol, or hydrogen. As these fuels ignite at lower temperatures, they will raise the cylinder temperature and thereby facilitate the ignition of ammonia.

However, the use of ammonia-fuel blends can lead to a reduced power output due to the nature of ammonia combustion. Yapicioglu and Dincer (2018) conducted a study to evaluate the effectiveness of different ammonia-fuel blends in a power generator engine. Their findings showed that as the ammonia content in the fuel mixture increased, the power output decreased. Particularly, the highest power output was achieved with an ammonia-hydrogen blend.

In addition to power output concerns, ammonia-fuel blends can impact CO_2 and NO_x emissions, depending on the complementary fuel used. The use of hydrogen in the ammonia-fuel blend suppresses the carbon emissions due to the absence of carbon in the fuel mixture. However, in both pure ammonia and ammonia-hydrogen blend combustion, NO_x emissions remain a significant concern (ERDEMIR; DINCER, 2021).

Because of the growing interest in these unconventional fuels, and due to the relevance of ignition delay time in practical application, a database encompassing relevant aspects of ammonia/ammonia hydrogen auto-ignition process defines the application case selected for this study.

2.2 IGNITION DELAY TIME

Ignition delay time (IDT) is generally defined as the time required for a fuel mixture to react and reach the stage of rapid heat release. This period is characterized by an increase in the concentration of radicals.

The IDT is dependent on temperature, pressure, and equivalence ratio, Φ (KHALED et al., 2017). To measure this delay, shock tubes (STs) and rapid compression machines (RCMs) are frequently used. ST experiments, which will be explained in detail in the next section, are widely used for high-temperature studies. On the other hand, RCM experiments are frequently employed as an alternative to STs for investigating low-to-intermediate temperature auto-ignition chemistry (SUNG; CURRAN, 2014).

The specific definition of IDT can vary depending on the experimental apparatus employed, with common methods relying on pressure, hydroxyl radical $OH^{\bullet}$ emission, or absorbance to infer the delay (LACKNER et al., 2013).

At the onset of ignition, the pressure and $OH^{\bullet}$ emission signals increase rapidly, whereas the absorbance signal declines rapidly. To determine the IDT, the steepest rise in the pressure and $OH^{\bullet}$ emission signals can be extrapolated backward to their respective baseline levels. Similarly, the fastest decay in absorbance signal, when extrapolated back to the pre-ignition absorbance level, can also be used to determine the IDT (CHOUDHARY et al., 2021)

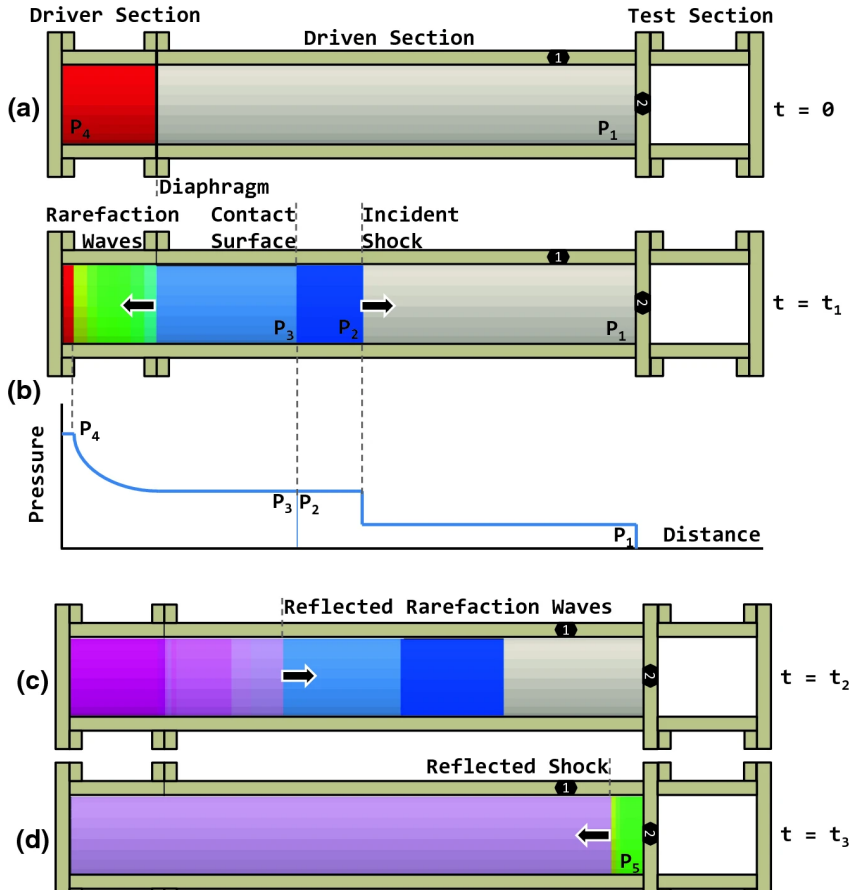
2.3 SHOCK TUBE

A shock tube experiment consists of a tube divided by a diaphragm into two sections: the driver section (high pressure P_4) and the driven section (low pressure P_1), as illustrated in Fig. 1a. When the diaphragm bursts due to the high pressure difference, a shock wave that propagates into the low-pressure section of the tube is

generated. The shock tube allows extensive control over pressure (0.1–1000 bar) and temperature (500–15000 K) with minimal effort (LACKNER et al., 2013).

The gas to be studied is placed in the driven section, while the driver section is filled with a highly compressed gas, usually hydrogen or helium. According to shock tube theory, the speed at which the shock wave propagates is independent of the driver section's volume (NISHIDA, 2001).

Figure 1 – Schematic illustration of events occurring inside the shock tube: (a) Initial conditions before the diaphragm rupture, (b) formation of various wavefronts immediately after the diaphragm bursts, with the pressure distribution shown along the tube, (c) propagation of reflected rarefaction waves into the driven section, and (d) shock wave upon reaching the end of the driven section.



Source: Source: Vivek and Sitharam (2020).

At time $t = 0$ (Fig. 1a), the diaphragm bursts, and the sudden release of the high-pressure driver gas into the low-pressure driven gas generates a series of compression waves that propagate through the driven section and raise its temperature.

By the time $t = t_1$ (Fig. 1b), the rear compression waves generate an incident shock wave that travels faster than sound in the driven gas. The driver and driven gases meet at a contact surface located just behind the shock wave, and the pressure between these regions is defined as P_2.

At the same time, a rarefaction wave moves in the opposite direction of the shock wave, traveling back into the driver section at the speed of sound. The pressure drop across the expansion wave is gradual, as illustrated by the pressure–distance curve in Fig. 1b. This region between the contact surface and the expansion wave is labeled P_3. Upon reaching the end of the driver section, the rarefaction head wave reflects and interacts with the incoming wave tail, further accelerating the reflected rarefaction wave (Fig. 1c) (VIVEK; SITHARAM, 2020).

Finally, at time $t = t_3$ (Fig. 1d), the incident shock wave reflects off the end wall of the driven section, causing an additional pressure increase, denoted as P_5. Typical test durations are 1 μs (LACKNER et al., 2013).

2.4 SHOCK TUBE SIMULATIONS

In addition to experimental methods, the ignition delay time can also be derived from combustion simulations. In this study, the simulation package FlameMaster (PITSCH, 1993) is used to obtain IDT data through these simulations.

The input parameters to the application align with those typically measured in ST experiments behind the reflected shock wave. A Newton solver is utilized to address the nonlinear system equations, described in the next subsection.

2.4.1 Governing Equations

For IDT simulations, the auto-ignition processes are modeled as a homogeneous, closed 0-D adiabatic, isochoric system, with prescribed non-reactive pressure profiles obtained from experimental studies in the literature. The simulation results are derived from a system of differential equations that model idealized configurations, approximating the conditions of shock tube experiments. Solving the governing equations for this configuration involves addressing an initial value problem, described by Warnatz et al. (2001):

$$\frac{dY_k}{dt} = \frac{\dot{m}_k}{\rho} \tag{1}$$

$$\frac{dT}{dt} = -\frac{1}{\rho c_V}\left(\sum_{k\in S}\left(h_k - \frac{R_u T}{M_k}\right)\dot{m}_k - \frac{P}{\rho}\frac{d\rho}{dt}\right) \tag{2}$$

$$Y_k(t_o) = Y_{k,o}, \quad k \in S \tag{3}$$

$$T(t_o) = T_0 \tag{4}$$

where ρ is the density, P the pressure, Y_k the mass fraction of the kth species in $S = [1, \ldots, K]$, T the temperature, R_u the universal gas constant, and c_V the specific

heat at constant volume. The chemical production rate $\dot{m}_k$ for species k is defined as the sum of contributions from all R reactions:

$$\dot{m}_k = \sum_{j=1}^{R} \dot{m}_{kj} \tag{5}$$

The chemical production rate for species k in reaction j, denoted $\dot{m}_{kj}$ is given by:

$$\dot{m}_{kj} = \omega \; M_k \; v_{kj} \tag{6}$$

where M_k represents the molecular weight of species k, ω is the reaction rate, explained with details in Sec. 2.5.2, and $v_{kj} = v''_{kj} - v'_{kj}$ is the difference between the molar stoichiometric coefficients of products (v''_{kj}) and reactants (v'_{kj}).

The temperature dependence of the species-specific heat at constant pressure is expressed as:

$$\frac{c_P(T)}{R_u} = a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4 \tag{7}$$

where a_1, a_2, a_3, a_4, and a_5 are the numerical coefficients supplied in NASA thermodynamic files (MCBRIDE, 2002). The relationship between the specific heats at constant volume and constant pressure is given by $c_P = c_V + R_u$.

The system of equations is completed with the ideal gas law $P = \rho(R_u/M)T$ where M is the mixture-averaged molecular weight.

FlameMaster requires three input files to process information about thermochemistry, transport properties, and the reaction steps governing the evolution of species involved in the combustion process. These files include thermodynamic data, transport data, and the reaction mechanism, which adheres to the CHEMKIN (REACTION DESIGN, 2016) format.

As the number of species in the kinetic model increases, the number of equations in the system that need to be solved also increases, which can lead to a significant increase in the computational cost. To address this, several reduction techniques have been developed over time. Section 2.6 provides an overview of the widely employed graph-based reduction approaches.

2.5 CHEMICAL KINETIC COMBUSTION MODELS

The Cambridge Dictionary (2024) defines combustion as a chemical process in which a fuel reacts with an oxidizing agent to produce heat and light. In other words, combustion transforms the energy stored in chemical bonds into heat that can be utilized in diverse energy applications, from power generation to transportation.

Understanding those interactions at a molecular level is important to the study of combustion, and chemical kinetic mechanisms serve as a structure to model them. The mechanism includes a set of species, each defined by its unique thermodynamic and transport properties, as well as a series of elementary chemical reaction steps that describe the interaction of this species in a combustion system.

In this section, key aspects of chemical kinetic combustion mechanisms are explored.

2.5.1 Global and Elementary Reactions

The combustion of one mole of fuel with A moles of oxidizer, producing B moles of combustion products, can be expressed by the global reaction mechanism (TURNS, 2000)

$$F + a\,Ox \rightarrow b\,Pr. \tag{8}$$

However, combustion is not a single-step process. Instead, it is a sequence of hundreds of elementary reactions where numerous chemical species take part. These elementary reactions describe interactions on a molecular level, and when combined, provide a detailed description of the combustion process. Most elementary reactions are bimolecular (TURNS, 2000), i.e. two molecules collide and react to form two new molecules. An arbitrary bimolecular reaction can be expressed as follows:

$$v_A\,\mathrm{A} + v_B\,\mathrm{B} \rightleftharpoons v_C\,\mathrm{C} + v_D\,\mathrm{D} \tag{9}$$

For instance, when the pure ammonia is ideally combusted, H_2O and N_2 are formed:

$$4\,\mathrm{NH_3} + 3\,\mathrm{O_2} \rightarrow 2\,\mathrm{N_2} + 6\,\mathrm{H_2O} \tag{10}$$

However, in reality, the process is more intricate. NUIG_2023 model for ammonia and ammonia/hydrogen combustion contains 306 reactions (ZHU et al., 2024). Even when considering only the most important elementary reactions during pyrolysis (GIRHE et al., 2024), the number of reactions involved increases significantly:

$$\mathrm{NH_3} + \mathrm{M} \rightleftharpoons \mathrm{NH_2} + \mathrm{H} + \mathrm{M} \tag{11}$$

$$\mathrm{NH_3} + \mathrm{H} \rightleftharpoons \mathrm{NH_2} + \mathrm{H_2} \tag{12}$$

$$\mathrm{NH_2} + \mathrm{M} \rightleftharpoons \mathrm{NH} + \mathrm{H} + \mathrm{M} \tag{13}$$

$$\mathrm{NH_2} + \mathrm{H} \rightleftharpoons \mathrm{NH} + \mathrm{H_2} \tag{14}$$

$$\mathrm{NH_2} + \mathrm{NH} \rightleftharpoons \mathrm{N_2H_2} + \mathrm{H} \tag{15}$$

$$\mathrm{NH_2 + NH_2 \rightleftharpoons NH_3 + NH} \tag{16}$$

$$\mathrm{NH_2 + NH_2 \rightleftharpoons N_2H_2 + H_2} \tag{17}$$

$$\mathrm{N_2H_2 + M \rightleftharpoons NNH + H + M} \tag{18}$$

where M represents third bodies, and is required to absorb the energy released in the formation of the stable species. During the collision, the internal energy of the newly formed molecule is transferred to a third body. If this transfer did not occur, the molecule would dissociate back into its original radicals due to the excess of internal energy (TURNS, 2000).

The set of elementary reactions required to represent an overall reaction is called a reaction mechanism (TURNS, 2000). The products of one reaction can serve as reactants in another. Therefore, a reaction mechanism can be understood as a series of elementary reactions. The goal of mechanism reduction is to identify the smallest number of elementary steps needed to accurately describe a particular global reaction.

2.5.2 Reaction Rates

The chemical reaction rates are dependent on temperature, pressure, and reactant concentration, and play a role in controlling the rate of combustion and determining the formation and destruction of pollutants (TURNS, 2000).

$$\omega = \omega_f - \omega_b = k_f \prod_{j=1}^{\mathsf{n_E}} [S_j]^{v'_j} - k_b \prod_{j=1}^{\mathsf{n_{Pr}}} [S_j]^{v''_j} \tag{19}$$

where the reaction rate constants k_f and k_b refer to the forward and backward rate coefficients, respectively. $\mathsf{n_E}$ and $\mathsf{n_{Pr}}$ denote the number of educts and products, while v'_j and v''_j represent the stoichiometric coefficients of species j on the reactant and product sides, respectively.

If the temperature range of interest is not too large, the bimolecular forward rate coefficient can be expressed by the Arrhenius empirical form,

$$k_f(T) = A \exp(-E_A/R_u T) \tag{20}$$

where R_u is the universal gas constant and T is the temperature. The exponential term expresses the fraction of collisions that occur with an energy exceeding the threshold required for the reaction, known as the activation energy E_A (LACKNER et al., 2013).

A is a constant named frequency factor, which represents the rate at which molecules collide. It is not strictly constant but depends on $T^{1/2}$ according to collision theory (TURNS, 2000). Therefore, the three-parameter functional form of the Arrhenius equation is frequently utilized:

$$k_f(T) = AT^b \exp(-E_A/R_u T) \tag{21}$$

where A, B, and E_A are the three empirical parameters. Arrhenius plots of $\log k$ versus $1/T$ for experimental data are used to determine the activation energy, as the slope of these plots is given by $-E_A/R_u$.

At equilibrium, the forward and backward reaction rates must be equal. Using this principle, the backward rate coefficient can be determined using the relationship:

$$\frac{k_f}{k_b} = K_C \tag{22}$$

At a process occurring at constant temperature and pressure, the correlation between the Gibbs free energy ΔG and the dimensionless equilibrium constant K'_C is (LACKNER et al., 2013):

$$K'_C = \exp(-\Delta G^0/R_u T) \tag{23}$$

If the enthalpies and entropies of the species at the given conditions are known, ΔG^0 can be calculated and K_C can then be obtained with the stoichiometric coefficients $v_i = v_a, v_b, \ldots$:

$$K_C = K'_C \left(\frac{P_{ref}}{R_u T}\right)^{\sum v_i} \tag{24}$$

where P_{ref}, usually 1 atm, is the pressure of reference.

Under certain conditions, reaction rate expressions depend on both pressure and temperature. For the high- and low-pressure limiting cases, the Arrhenius rate coefficients, k_∞ and k_0 are calculated, respectively (REACTION DESIGN, 2016). The rate coefficient at any given pressure is then determined by:

$$k = k_\infty \left(\frac{P_r}{1 + P_r}\right) \tag{25}$$

where the reduced pressure P_r is given by

$$P_r = \frac{k_0\, C}{k_\infty} \tag{26}$$

and C represents the total concentration of the mixture.

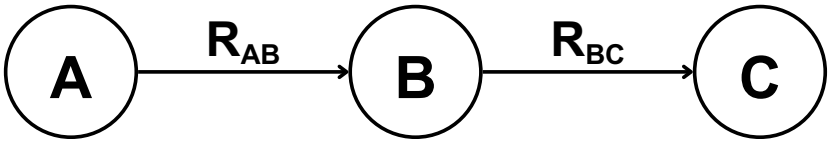
2.6 GRAPH-BASED REDUCTION TECHNIQUES

The goal of the reduction procedure is to identify a set of species in the detailed mechanism that can be removed with minimal impact on the target species. This is achieved by assigning direct interaction coefficients to each species, which quantify

how strongly each species interacts with others. Species with low interaction coefficients can be safely removed. As a result, a skeletal mechanism is created by removing any reactions from the detailed mechanism that involve the removed species as either reactants or products.

The reaction network can be converted into a directed weighted graph $G = \{e, v, w\}$, where the vertices (v) represent the species, edges (e) the reactions, and the edge weights (w) indicate the interaction strength between species. As illustrated in Fig. 2, the reactions are shown as arrows: one arrow indicates the conversion of species A to B, and the next represents the conversion of species B to C. The coefficient R_{AB}, which varies depending on the reduction method, quantifies the interaction strength between species A and B and will be defined in the following sections.

Figure 2 – Reaction network graph.



Source: Author (2024).

This approach is simple and fast, requiring minimal user involvement in the reduction process. The following subsections outline various existing graph-based reduction techniques and their methods for quantifying the degree of interaction between species.

2.6.1 DRG

In the Directed Relation Graph (DRG) method (LU; LAW, 2006), the weight of the directed edge linking a species A to species B is proportional to the contribution of B in the production rate of A, estimated as

$$R_{\mathsf{AB}}^{DRG} \equiv \frac{\sum_{i=1}^{n_R} |v_{i,A}\omega_i\delta_B^i|}{\sum_{i=1}^{n_R} |v_{i,A}\omega_i|} \tag{27}$$

where n_R denotes the total number of reversible reactions. For each reaction i, the forward, backward, and net reaction rates are represented by $\omega_{f,i}$, $\omega_{b,i}$, and ω_i, respectively. δ_B^i is defined as

$$\begin{cases} \delta_B^i = 1 & \text{, if the } i\text{th reaction involves species B} \\ 0 & \text{, otherwise} \end{cases} \tag{28}$$

If A and B are directly related, then both R_{AB} and R_{BA} are nonzero and typically not equal to each other. In a graph, R_{AB}, representing the influence of species B on species A, will be represented as a directed edge from A to B.

The definition given in Eq. 27 represents the error involved in predicting species A when species B is excluded. After calculating all the coefficients, a user-defined cutoff threshold ϵ is applied, and edges where $R_{\mathsf{AB}} < \epsilon$ are removed. Next, a search is performed starting from the target species, where only those species that can be reached are kept. Any reactions involving species that are not included in this search are removed.

With this approach, only one evaluation using the detailed mechanism is needed. However, it assumes that all species kept in the mechanism are equally important and that all strongly linked species need to be kept, which might not always be necessary.

The DRG method treats production and consumption reactions the same way, but removing a species that only affects the consumption of target A will not impact A's net production rate as much as removing a species that affects both production and consumption equally. For the same R_{AB} value, the first case will cause a bigger error in the net production rate of A than the second case, where the error might be minimal because removing it also removes some of the consumption effects (PEPIOT-DESJARDINS; PITSCH, 2008).

2.6.2 DRGEP

With the assumption that species being far from the specified target are less important, the Directed Relation Graph with Error Propagation (DRGEP) (PEPIOT-DESJARDINS; PITSCH, 2008) is defined, considering the net contribution of species B to species A.

In the DRGEP formulation, the set of species to be retained in a simplified kinetic scheme is determined by calculating condition-specific coefficients, R_{AB}, which provide an a priori estimate of the error in predicting a target species A when a species B is removed. These coefficients are defined as:

$$R_{\mathsf{AB}} = \max_{\text{all paths p}} (r_{\mathsf{AB,\,p}}) = \max_{\text{all paths p}} (f_{\mathsf{p}}(P, T, \boldsymbol{Y})) \tag{29}$$

where

$$r_{\mathsf{AB,\,p}} = \prod_{k=1}^{n-1} r_{S_k S_{k+1}} \tag{30}$$

is the path-dependent interaction coefficient between target species A and species B, where p identifies all the possible pathways connecting the two species.

The coupling strength between two species S_k and S_{k+1} can be expressed as

$$r_{S_k, S_{k+1}} \equiv \frac{\left| \sum_{i=1}^{n_R} v_{i,k} \omega_i \delta^i_{S_{k+1},\{S\}} \right|}{\max(P_k, C_k)} \tag{31}$$

where n_R denotes the total number of reversible reactions and $\{S\}$ is the set of species already removed. $v_{i,k}$ is the stoichiometric coefficient of species k in the ith reaction.

$\delta^{i}_{S_{k+1,\{S\}}}$ is unity if the ith reaction involves species S_{k+1} or any species in subset {S}, and 0 otherwise. The production and consumption rates are calculated as follows:

$$P_k = \sum_{i=1}^{n_R} \max\left(0, v_{i,k}\,\omega_i\right) \tag{32}$$

$$C_k = \sum_{i=1}^{n_R} \max\left(0, -v_{i,k}\,\omega_i\right) \tag{33}$$

This way, if some error is introduced in the prediction of species B, the longer the way this error has to propagate to reach target A, the smaller its effect will be, typically.

Compared to the DRG (LU; LAW, 2006), the normalization factor in Eq. 31 allows to account for the influence of the relative contributions of a species' consumption or production on the adjacent species, allowing to assign a minor weight to those interactions where a species contribute to the consumption of the adjacent one and the latter is mostly produced or vice-versa.

However, for the indirect relations, DRGEP only identifies the strongest reaction pathway, making it inadequate for accurately representing species flux when multiple intermediate species are present in parallel (SUN et al., 2010). Furthermore, the interaction coefficient fails in describing relations between species with both high production and consumption rates, such as those with catalytic behavior (SUN et al., 2010).

2.6.3 PFA

All the previously mentioned methods are based on the analysis of one-generation reaction rates. The path flux analysis (PFA) (SUN et al., 2010) method analyzes the formation and consumption fluxes of each species across multiple reaction path generations, identifying the important reaction pathways and the associated species. The important reaction pathways are identified with the production and consumption fluxes, defined in Eq. 32 and 33, respectively.

To account for the conservative flux information, a modified definition is introduced, containing both the first and second generations. Although the method can be extended to any generation, the computation time increases exponentially.

The interaction coefficients for the production and consumption of species A due to species B in the first generation are defined as:

$$R^{pro-1st}_{\mathsf{AB}} = \frac{P_{AB}}{\max(P_A, C_A)} \tag{34}$$

$$R^{con-1st}_{\mathsf{AB}} = \frac{C_{AB}}{\max(P_A, C_A)} \tag{35}$$

Using Eq. 34 and 35, the interaction coefficients, which represent the ratios of fluxes between species A and B via a third reactant (species M_i) for the second generation, are defined as:

$$R_{\mathsf{AB}}^{pro-2nd} = \sum_{M_i \neq A,B} \left(R_{AM_i}^{pro-1st} \, R_{M_iB}^{pro-1st} \right) \tag{36}$$

$$R_{\mathsf{AB}}^{con-2nd} = \sum_{M_i \neq A,B} \left(R_{AM_i}^{con-1st} \, R_{M_iB}^{con-1st} \right) \tag{37}$$

with the threshold being defined as the sum of all the interaction coefficients, so that if $R_{\mathsf{AB}} \equiv R_{\mathsf{AB}}^{pro-1st} + R_{\mathsf{AB}}^{con-1st} + R_{\mathsf{AB}}^{pro-2nd} + R_{\mathsf{AB}}^{con-2nd} > \epsilon$, species B is added to the set of important species.

Based on the indexes of each reaction path for the first and second generations, reduced chemical mechanisms of varying sizes can be generated.

3 METHODOLOGY

This section provides an overview of the framework implemented for the clustering and reduction procedure. The flowchart is presented in Fig. 3.

The main assumption of this study is that if the detailed model-derived reaction networks, describing how a combustion system evolves given an initial thermodynamic state, are similar, then the simplified kinetic schemes valid for those states will also exhibit similarity. This assumption implies that when numerical solutions under different experimental conditions exhibit comparable reaction networks, a simplified model for one of those solutions can be reliably applied to others with similar characteristics.

To evaluate similarities between different experimental conditions, a similarity metric based on the comparison of reaction networks is introduced. Each experiment in the dataset is represented by a reaction network modeled as a weighted directed graph, following the approach employed in the DRGEP method (PEPIOT-DESJARDINS; PITSCH, 2008). Each case is then processed through the ranking algorithm developed by Wang et al. (2024) in the species ranking step, which converts the condition-specific graph into a probability distribution that reflects the interconnections among species and their structural roles within the reaction network.

The assumption enables the reduction of detailed models more efficiently by systematically identifying and exploiting similarities among reaction networks. The complete set of conditions representing the applicability domain of the detailed kinetic model can be grouped into distinct clusters of similar conditions in the affinity propagation step. Within each cluster, a representative state can be then selected to be used for model reduction.

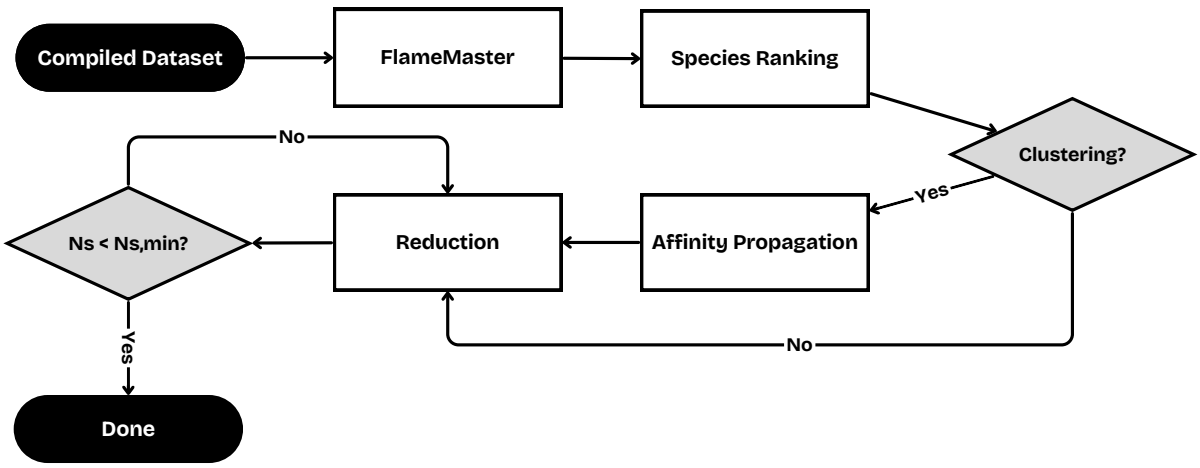
The similarity between the probability distributions of different conditions is quantified using the Jensen-Shannon Divergence (JSD), and serves as input for a clustering algorithm based on affinity propagation (FREY; DUECK, 2007).

Following clustering, the reduction process is performed for the representative cases identified within each cluster. If the minimum number of species $N_{s,min}$ is reached, the process is done. The effectiveness of the reduced models is assessed using a curve-matching score to ensure accuracy to the original detailed models.

3.1 DATASET

The combustion simulations were carried out using the FlameMaster software package (PITSCH, 1993). A dataset comprising 264 ignition delay time experiments for ammonia combustion was used, covering a broad range of conditions, including temperatures, pressures, equivalence ratios (Φ), H_2-blending ratios, and dilutions. A

Figure 3 – Flowchart of the implemented framework.



Source: Author (2024)

summary of the dataset is provided in Table 1.

Table 1 – Summary of the experimental data.

Author	Mixture	T [K]	P [bar]	Φ	Data points	Reference
B.Shu	$NH_3/N_2/O_2$	1181 - 1581	20, 40	0.5, 1, 2	30	Shu et al. (2019)
J.Chen_2021	$NH_3/O_2/H_2/Ar$	1022 - 1957	1.2, 10	1	66	Chen et al. (2021)
J.Baker_2023	$NH_3/N_2/O_2/H_2/Ar$	1511- 1774	2,2	1	5	Baker et al. (2022)
O.Mathieu	$NH_3/O_2/Ar$	1564 - 2489	1.4, 11, 30	0.5, 1, 2	103	Mathieu and Petersen (2015)
Peng_2023	$NH_3/O_2/Ar$	1180 - 1941	10 - 22.4	0.5, 1, 2	60	Peng et al. (2023)

Source: Author (2024)

The data set was compiled by using XML files collected from each experimental data, containing details about initial temperature, initial pressure, initial composition, and the IDT definitions. The IDT definitions were consistently applied based on those established in the experimental studies documented in the literature.

The files were processed into appropriate inputs for the simulation software, providing the initial conditions for the equations 1–4. The equations were then discretized into a nonlinear system of algebraic equations and integrated over time using a variable-order, stiff-optimized time-stepping algorithm implemented in the CVODE package (COHEN et al., 1996).

Although various chemical kinetic models are available, their validation often shows strong performance only within a particular range of conditions. Girhe et al. (GIRHE et al., 2024) performed a quantitative evaluation of 16 recent kinetic models for pure NH_3 and NH_3/H_2 mixture combustion. The evaluation employed a similarity score, derived by comparing the experimental data with their corresponding model predictions. Among the evaluated models, the NUIG_2023 mechanism exhibited the best overall agreement with experimental IDT measurements.

Therefore, the simulations in this study were conducted using the NUIG_2023 chemical kinetic model (ZHU et al., 2024), which comprises 43 species and 306 reactions.

3.2 REACTION NETWORK

In order to perform the evaluation of the reaction network, an adjacency matrix was constructed to represent the interactions among species based on their reaction fluxes.

The matrix $\mathbf{A}$ is a square matrix where each row and column corresponds to a specific species. Each entry in the matrix quantifies the connection between a pair of species, with zero values indicating no connection. The coupling strength is computed using Eq. 31, which measures the production and consumption fluxes for all reactions a species participates in. This matrix serves as the main input to the following ranking routine.

3.3 SPECIES RANKING

To evaluate the importance of each species within the reaction network, the method recently proposed by Wang et al. (2024), referred to as Directed Relation Graph-based SpeciesRank (DRGSR), was employed in this work. The method modifies the formulation of the PageRank (BRIN; PAGE, 1998) algorithm to incorporate information about the weight in a directed-weighted graph. Details about the method are provided below.

For each node in a graph, the PageRank algorithm calculates the relative importance of every node pointing to it and assigns a PageRank value accordingly. A node can have a high PageRank if many nodes point to it, or if the few nodes that point to it have a high PageRank value.

Let u be a node. Then let F_u be the set of nodes u points to, and B_u be the set of nodes that point to u. Let $N_u = |F_u|$ be the number of edges from u. A simple ranking pr can be defined as follows:

$$pr_U = \sum_{v \in B_u} \frac{R(v)}{N_v} \tag{38}$$

In matrix notation, let $\mathbf{A}$ be a square matrix with the rows and columns corresponding to nodes. Let $A_{u,v} = 1/N_u$ if there is an edge from u to v and $A_{u,v} = 0$ if not. If $\mathbf{pr}$ is treated as a vector over nodes, then:

$$\mathbf{pr}_{i+1} = \mathbf{A}^T \mathbf{pr}_i \tag{39}$$

However, during the iteration process to compute the PageRank, if there are nodes that point to each other but to no other node, and a node that points to one of them, this loop will accumulate rank but never distribute any rank. To address this issue, the damping factor d is introduced (PAGE et al., 1998):

$$\mathbf{pr}_{i+1} = (d)\,\mathbf{A}^T\mathbf{pr}_i + (1-d)\,\mathbf{1} \tag{40}$$

where $\mathbf{1}$ is the vector consisting of all ones.

When nodes have no outgoing edges, it is unclear where their weight should be distributed. Since those nodes do not influence the ranking of other nodes, they can be resolved by adding edges from the dangling nodes to all nodes, without significantly affecting the overall ranking (THALER, 2013). The formula is modified as follows:

$$\mathbf{pr}_{i+1} = (d)\left(\mathbf{A}^T + \mathbf{1}\times\mathbf{e}^T\right)\mathbf{pr}_i + (1-d)\,\mathbf{1} \tag{41}$$

where vector $\mathbf{e}$ represents existence of dangling nodes,

$$\begin{cases} \mathbf{e}_i = 1/n & \text{, if } N_i = 0 \\ 0 & \text{, otherwise} \end{cases}$$

being n the total number of nodes in the graph.

The entries in $\mathbf{A}^T + \mathbf{1}\times\mathbf{e}^T$ are all non-negative, and the sum of the entries in a column is one. Matrices with those configurations are called stochastic, and every column-stochastic matrix has 1 as an eigenvalue.

Therefore, the ranking equation is essentially a standard problem of finding an eigenvector for a square matrix. The eigenvalues λ and eigenvectors $\mathbf{x}$ of a matrix $\mathbf{A}$ satisfy the equation $\mathbf{Ax} = \lambda\mathbf{x}, \mathbf{x} \neq 0$. We thus seek an eigenvector $\mathbf{x}$ with eigenvalue 1 for the matrix $\mathbf{A}$.

During the ranking phase, a damping factor of 0.85 was applied, while species that are not connected to any nodes were assigned a rank of 0. Wang et al. (2024) modified the PageRank formulation to account for the weighted graphs. Hence, in their formulation, the matrix $\mathbf{A}$ is defined as described in Section 3.2.

For each case in the study, the ranking routine returns a ranking vector, denoted $\mathbf{pr}$, which is an N_s-dimensional vector, where N_s represents the total number of species in the model. This vector represents the discretized probability distribution that characterizes the likelihood of the combustion system evolving toward specific nodes.

In simple terms, the higher the value of a component, the greater the influence of the corresponding species within the network of reactions for that specific case. This allows for an assessment of how critical each species is in the context of the combustion process.

Once the ranking vector $\mathbf{pr}$ is determined for all considered cases, the user has the option to proceed with the clustering. If clustering is not enabled, the final ranking assigned to a species s_k, SR_{s_k} is computed as:

$$SR_{s_k} = \max_{t \in N}(pr_{s_k,t}), \tag{42}$$

where N is the total number of cases considered in the database of numerically computed solutions. If the clustering option is enabled, the cases in the database are grouped based on the affinity propagation algorithm.

3.4 CLUSTERING

The clustering procedure was performed using the affinity propagation algorithm from the Scikit-learn Python package (PEDREGOSA et al., 2011). A detailed explanation of how the algorithm works is provided in the next section.

3.4.1 Affinity propagation

Affinity propagation (FREY; DUECK, 2007) is a clustering method based on the concept of passing messages between data points. Unlike other methods, which require the number of clusters to be determined before running the algorithm, affinity propagation considers all data points as potential cluster centers, called exemplars.

The method takes as input measures of similarity between pairs of data points, where the similarity $s(i,k)$ indicates how well the data point k is suited to be the exemplar for data point i. The self-similarity value, $s(k,k)$, is referred to as "preference". Data points with larger values of preference are more likely to become an exemplar.

The affinity propagation algorithm treats each data point as a node in a network, where messages are recursively passed along the edges between nodes until a good set of exemplars and their corresponding clusters is identified.

There are two types of messages being passed. The “responsibility” $r(i,k)$, sent from point i to candidate exemplar point k, quantifies how well-suited point k is to serve as the exemplar for point i, relative to other potential exemplars for point i:

$$r(i,k) \leftarrow s(i,k) - \max_{k' \neq k}\left(a(i,k') + s(i,k')\right) \tag{43}$$

The “availability” $a(i,k)$, sent from candidate exemplar point k to point i, quantifies how appropriate it would be for point i to choose point k as its exemplar, taking into account the other points’ preference for k as an exemplar:

$$a(i,k) \leftarrow \min\left(0, r(k,k) + \sum_{i' \notin \{i,k\}} \max\left(0, r(i',k)\right)\right) \tag{44}$$

Both responsibility and availability matrices are initialized to zero. At each iteration, the algorithm then updates these matrices and identifies which points are exemplars. The process alternates between updating the responsibility and availability matrices until convergence is reached, i.e. when messages stabilize and show minimal changes between iterations. Once the exemplars are determined, each data point is assigned to the cluster associated with its chosen exemplar.

The inputs to the algorithm are:

- $\mathbf{S}$: $N \times N$ matrix whose elements $s_{i,j}$ quantify the similarity between the reaction network associated to case i and j, respectively.
- $\mathbf{p}$: N-vector whose components denote the preference value assigned to a single case.

3.4.1.1 Similarity Matrix

With affinity propagation, the similarity between data points can be quantified in different ways, e.g. using the negative squared distance between them. In this study, the similarity is quantified by comparing case-specific ranking vectors, which represent the importance of species in the reaction network. The ranking assigned to species constitutes the elements of a discrete probability distribution. By measuring how those probability distributions are similar, the reaction network can be encoded into a quantifiable representation.

The distance between probability distributions can be quantified through the Kullback-Leibler Divergence (KLD) (KULLBACK; LEIBLER, 1951), which measures how one probability distribution O is different from a second probability distribution N. It is defined as:

$$\mathsf{KL}(O||M) = \sum_i O(i) \log \frac{O(i)}{M(i)} \tag{45}$$

In general, the KLD is an asymmetric distance (i.e., $\mathsf{KL}(O||N) \neq \mathsf{KL}(N||O)$. To address this limitation, the symmetrization of the KLD may be obtained by the Jensen–Shannon divergence (JSD) (NIELSEN, 2019). The JSD can be interpreted as the total KLD to the average distribution, and it is defined as:

$$\mathsf{JSD}(O||N) = \frac{1}{2}\mathsf{KL}(O||M) + \frac{1}{2}\mathsf{KL}(N||M) \tag{46}$$

where $M = 0.5(O + N)$ is the average distribution between O and N.

As JSD values approach zero, the similarity between the distributions increases. Therefore, the matrix element $s_{i,j}$, representing the similarity between cases i and j, is defined as:

$$s_{i,j} = 1 - \mathsf{JSD}(\mathbf{pr}_i||\mathbf{pr}_j) \tag{47}$$

3.4.1.2 Preferences

A uniform preference value was assigned to all cases, corresponding to the median of the similarity matrix (Eq. 48). This ensured that all cases were equally considered as potential cluster centers.

$$s(k,k) = \text{median}(\mathbf{S}) \tag{48}$$

This value resulted in a sufficiently small number of clusters, with their centers effectively capturing the differences in the underlying chemistry that characterize various auto-ignition regimes, as demonstrated in the following sections.

3.5 REDUCTION

Once the clusters and their corresponding centers (C) are identified, the final ranking associated with each species is computed as

$$SR_{s_k} = \max_{t \in C}(pr_{s_k,t}) \tag{49}$$

Based on their SR values, species are sorted in ascending order. The reduction process is then carried out by sequentially removing species until a user-specified minimum number is reached, $N_{s,min}$. At each reduction step, the accuracy of the reduced model is assessed using the Curve Matching score. This approach allows the performance of the reduced model to be compared against both experimental data and the results obtained with the detailed model. A detailed explanation of the comparison metric used for assessing the reduced models' performance is provided in the subsequent section.

3.6 EVALUATION

The Curve Matching (CM) (RAMALLI et al., 2023) method quantifies the similarity between two curves with a normalized score bounded between 0 and 1, where 1 indicates perfect similarity.

Consider two continuous curves f and g, obtained through an interpolation approach, with their respective derivatives, denoted as f' and g'. Let D be the intersection of the domains of f and g. The norm of a generic curve h in the L^2 space is defined as:

$$||h|| = \sqrt{\int_D h(x)^2 \, dx} \tag{50}$$

This way, the computation of the following four dissimilarity measures, which compose the CM score, can be done:

1.
$$d^0_{L^2}(f,g) = \frac{1}{1+\frac{||f-g||}{\sqrt{D}}} \in [0,1] \tag{51}$$

The first dissimilarity computes the L^2-norm of the difference between the two functions, being a generalization of the Sum Squared Error (SSE) to the continuous case.

2.
$$d^1_{L^2}(f,g) = \frac{1}{1+\frac{||f'-g'||}{\sqrt{D}}} \in [0,1] \tag{52}$$

The second dissimilarity extends the SSE to the first derivatives. If $f(x) = g(x)+k$, with $k \in \mathbb{R}$, then $d^1_{L^2} = 0$. Therefore, $d^1_{L^2}$ is invariant to vertical translations, but quantifies quantifies the similarity of slopes between the functions.

3.
$$d^0_{Pe}(f,g) = 1 - \frac{1}{2}\sqrt{\left|\left|\frac{f}{||f||} - \frac{g}{||g||}\right|\right|} \in [0,1] \tag{53}$$

The third dissimilarity is the Pearson correlation index, and measures whether the trend of a function is in agreement or not with the other. If $f(x) = g(x) \times k + a$, with $k, a \in \mathbb{R}$, then $d^0_{Pe} = 1$. The Pearson index is invariant to translation and dilation.

4.
$$d^1_{Pe}(f,g) = 1 - \frac{1}{2}\sqrt{\left|\left|\frac{f'}{||f'||} - \frac{g'}{||g'||}\right|\right|} \in [0,1] \tag{54}$$

The last dissimilarity is the Pearson correlation index applied to the first derivatives.

Fig.4 shows some characteristic cases where the four indices are quantified.

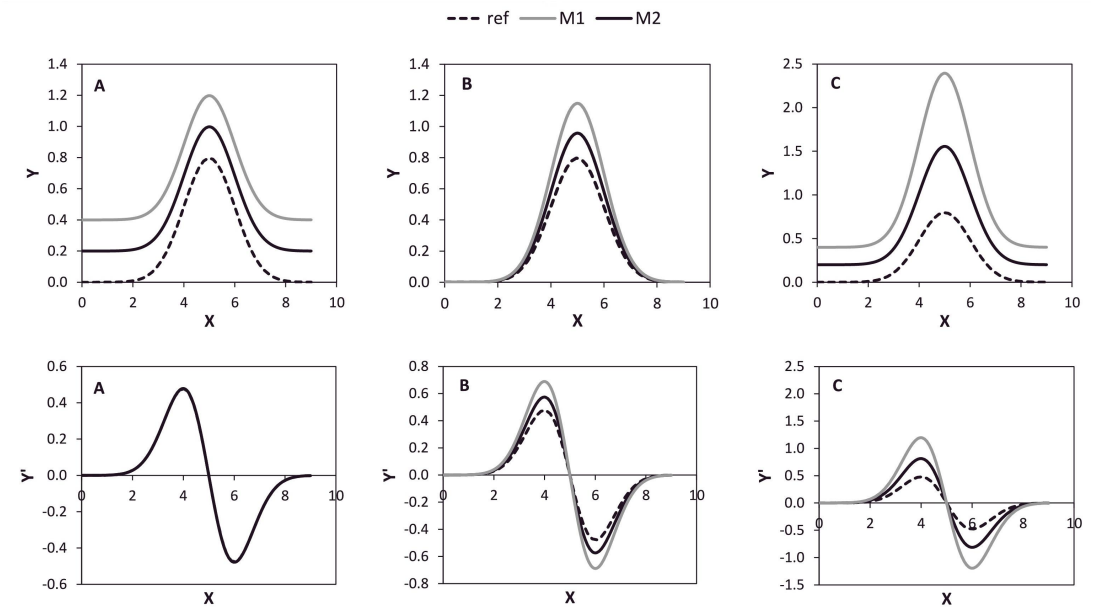
Compared to the original formulation by Ramalli et al. (RAMALLI et al., 2023), a square root transformation was applied. In the first two dissimilarities, it ensures non-dimensionality, allowing quantitative comparison across variables with different domain sizes. In the last two dissimilarities, the square root transformation establishes a linear relationship with the Pearson correlation coefficients.

The CM score for the IDT data is calculated as the arithmetic average of the four dissimilarities:

$$\text{CM}(f,g) = \frac{d^0_{L^2} + d^1_{L^2} + d^0_{Pe} + d^1_{Pe}}{4} \tag{55}$$

In its original formulation, a shift index S is also included to account for horizontal translations between the two functions. As the curve differences in this study are limited to the vertical direction, the CM score is computed using only the four non-shifted dissimilarities.

Figure 4 – Panel A: The curves differ by a vertical translation, panel B: The curves differ by a vertical dilation, panel C: The curves differ by a linear affine transformation of the y-axis. Respective dissimilarity values are reported below each panel.



	$d^0_{L^2}$	$d^1_{L^2}$	d^0_P	d^1_P		$d^0_{L^2}$	$d^1_{L^2}$	d^0_P	d^1_P		$d^0_{L^2}$	$d^1_{L^2}$	d^0_P	d^1_P
M_1	0.067	0.000	0.050	0.000	M_1	0.023	0.016	0.000	0.000	M_1	0.135	0.058	0.023	0.000
M_2	0.133	0.000	0.113	0.000	M_2	0.059	0.041	0.000	0.000	M_2	0.280	0.124	0.036	0.000

Source: Bernardi et al. (2016)

The CM score was used to quantitatively assess the performance of both reduced and detailed models.

For the evaluation of detailed models, the experimental data was used as reference. Experimental uncertainty is incorporated by performing bootstrap resampling (EFRON; TIBSHIRANI, 1994) around the experimental points within their uncertainty range and using the mean value of the resulting probability distribution as the comparison term.

When assessing reduced models, their performance can be evaluated either against the detailed model results or directly against experimental data. In the first case, the CM scores quantify how the performance of the reduced model deviates from that of the detailed model, and no uncertainty is applied. In the second case, where experimental data serve as the reference, it is important to note that comparisons may sometimes yield misleading results.

For instance, the CM score of the reduced model might be higher than that of the detailed model. While this might seem promising at first glance, it often indicates a compromise in the accuracy of the underlying chemistry. To avoid such ambiguities, this study focuses on comparing the performance of reduced models solely against the

detailed model.

4 RESULTS AND DISCUSSION

In this chapter, the results of the study are presented and discussed, focusing on the evaluation and validation of the methods used for the clustering and skeletal reduction. Section 4.1 provides general considerations, beginning with a comparison between the DRGEP and DRGSR models, followed by an evaluation of the clustering algorithm. Section 4.2 assesses the selection of centers, discussing their effectiveness and relevance. Finally, Section 4.3 evaluates the performance of the reduction mechanism based on the validation targets and based on the entire dataset.

4.1 GENERAL CONSIDERATIONS

In this section, a comparison between the DRGEP (PEPIOT-DESJARDINS; PITSCH, 2008) method and DRGSR (WANG et al., 2024) is shown to illustrate the motivation behind selecting the DRGSR method as a means to obtain a quantifiable representation of the reaction network. A simple example is then presented to demonstrate the validity of the clustering algorithm. Furthermore, a comparison with Principal Component Analysis (PCA) (GOKULAKRISHNAN et al., 2006) clustering is used to illustrate the reasons for favoring a clustering approach based on reaction flux comparisons over one that relies on the thermodynamic properties commonly defining the initial state of a combustion system.

4.1.1 Comparison of DRGEP and DRGSR

While the DRGEP approach has been widely employed to simplify complex reaction schemes, certain aspects of the method need closer examination. As an example, a reaction network defined exclusively by three chemical pathways is represented in the graph shown in Fig. 5.

Assuming node 0 as the target, the corresponding DRGEP coefficients associated with each species are presented in Table 2.

Nodes 4 and 5 present an interesting comparison: Node 4 is connected to species 1 and 2, whose DRGEP coefficients relative to the target node 0 ($R_{0,1}$, $R_{0,2}$) are identical. In contrast, node 5 is connected exclusively to node 3. Yet, due to a slight variation in the direct interaction coefficients, specifically between species 5 ($r_{3,5}$) and species 4 ($r_{2,4}$, $r_{1,4}$), species 5 is ranked higher than species 4. As a result, node 4 will be removed before node 5, even though its removal affects two of the three pathways.

This example highlights a limitation of the DRGEP method, which fails to capture the topological information of the reaction network, neglecting the specific role each species plays within it.

Figure 5 – Weighted graph.

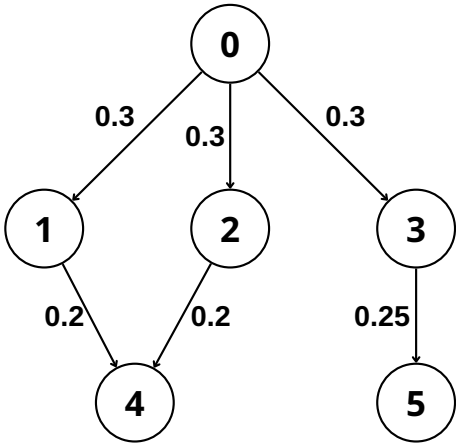


Table 2 – DRGEP coefficients assuming species 0 as target species.

Node	DRGEP
0	1
1	0.3
2	0.3
3	0.3
4	**0.06**
5	**0.075**

Source: Author (2024).

The method employed by Wang et al. (2024) addresses this limitation, as the algorithm prioritizes local information rather than relying solely on a global pathway coefficient. A comparison between the rankings generated by DRGEP and the new approach is presented in Table 3.

Table 3 – Comparison between DRGEP and Novel method rank scores on a weighted graph.

Node	DRGEP	DRGSR
0	1	0.1
1	0.3	0.13
2	0.3	0.13
3	0.3	0.13
4	**0.06**	**0.31**
5	**0.075**	**0.20**

Source: Author (2024).

4.1.2 Evaluation of the Clustering Algorithm

To assess the effectiveness of the clustering approach, a test was performed using nine cases with identical composition, dilution, pressure, and equivalence ratio, but differing in temperature. The temperatures ranged from 1050 K to 2000 K, representing low, intermediate, and high-temperature regimes. The conditions are detailed in Table 4.

The pairwise similarities between each case, based on the ranking of the reaction rate structures, are shown in Fig. 6a. These similarity values were used as input for the affinity propagation clustering algorithm.

The algorithm effectively identified distinct regions corresponding to the various temperature regimes. The scatter plot on Fig. 6b highlights the identified clusters,

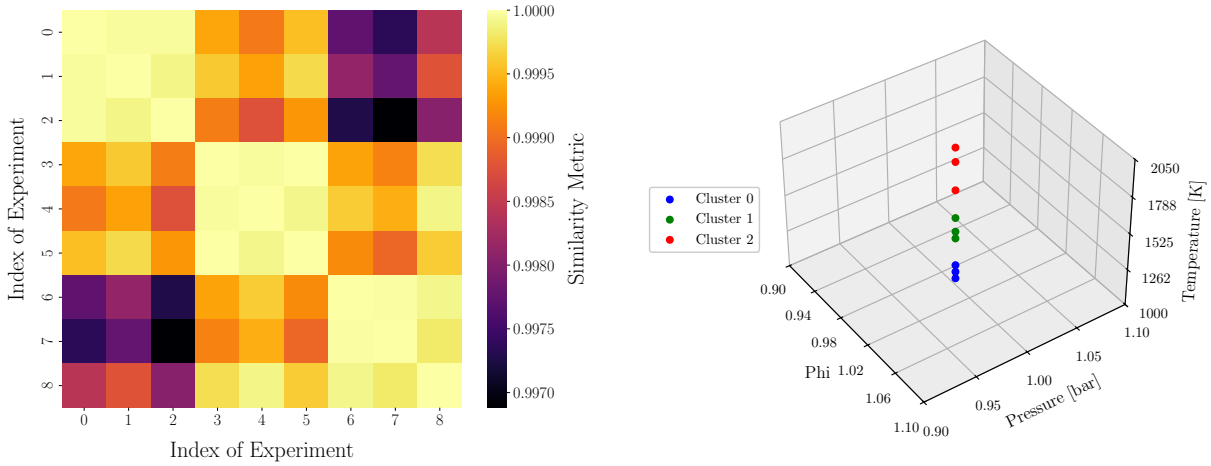
Table 4 – Experimental conditions for the assessment of the clustering.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	Dilution
●	0	1100.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	1	1150.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	2	1050.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	3	1400.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	4	1500.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	5	1350.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	6	1900.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	7	2000.0	1.0	1.0	$NH_3/N_2/O_2$	3.76
●	8	1700.0	1.0	1.0	$NH_3/N_2/O_2$	3.76

Source: Author (2024).

represented in a 3D plot of pressure, temperature, and Φ, where experiments 0, 3, and 6 were assigned as exemplars. This demonstrates the algorithm's ability to distinguish different thermodynamic regimes based on similarities in their reaction network structures.

Figure 6 – Visualization of the validation case: (a) Similarity matrix, and (b) scatter plot of initial pressure, temperature, and equivalence ratio highlighting cluster distributions.



(a) (b)

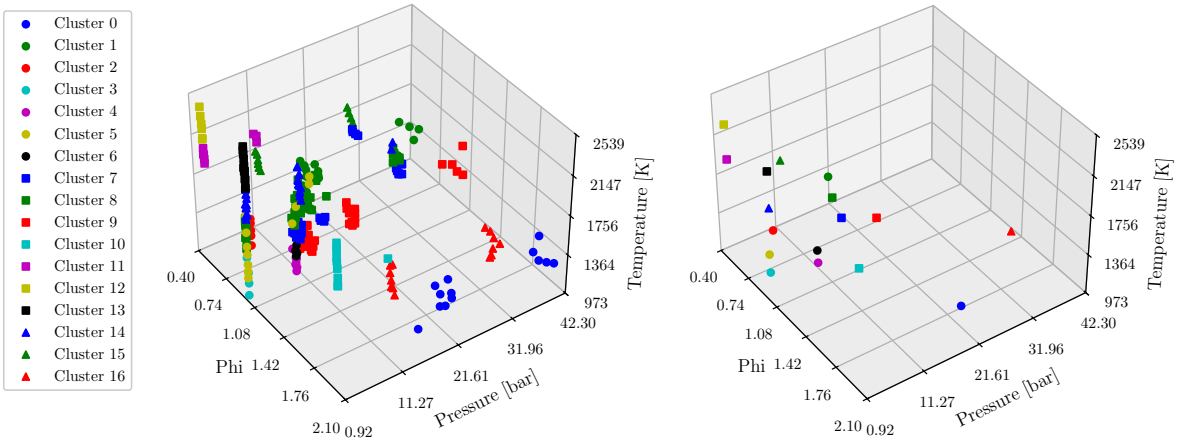
Source: Author (2024).

4.1.3 Affinity Propagation and Principal Component Analysis

The clustering algorithm was applied to the dataset of 264 cases, which resulted in the identification of 17 clusters. The distribution of these clusters can be visualized in Fig. 7, where the cases are plotted against pressure, temperature, and equivalence ratio.

It can be observed that different pressure, equivalence ratio and temperature values can lead to the same cluster, due to them having similar reaction fluxes. The

Figure 7 – Distribution of the 264 cases across 17 identified clusters: (a) all points, and (b) cluster centers.



(a) (b)

Source: Author (2024).

inverse is also true, where some points with the same initial properties exhibit distinct reaction fluxes.

To better visualize the clusters, Principal Component Analysis (PCA) (GOKULAKRISHNAN et al., 2006) was applied. PCA is a technique used to reduce the dimensionality of data. It transforms the original data set into a new set of variables called principal components (PCs). Those components are linear combinations of the original variables, computed in a way to retain as much of the original variance across the data as possible.

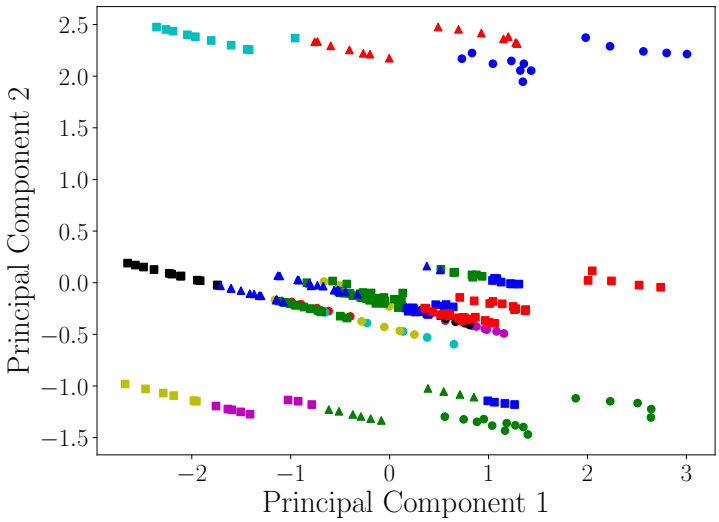
The principal components are ranked by their eigenvalues in descending order, so the first component captures the maximum possible variance. Components with lesser importance can be ignored when doing the transformation, so the final set will have fewer dimensions than the original.

The choice to reduce the data from three to two dimensions was justified by the fact that PC1 and PC2 together explain 80% of the total variance, a significant portion that retains most of the important information. The 3D data projected onto both the first and second components can be seen in Fig.8.

For equivalence ratios other than unity, a correlation between temperature and the resulting cluster is apparent. However, in experiments where the equivalence ratio is 1, the clusters overlap considerably. Therefore, while initial conditions such as pressure, temperature, and equivalence ratio are critical in defining the combustion process, they do not fully account for its evolution.

If clustering were to be performed solely using temperature, equivalence ratio, and pressure as features, the resulting cluster would fail to account for the differences

Figure 8 – PC1 and PC2.



Source: Author (2024).

in reaction fluxes and network structures. This approach could lead to grouping states that evolve very differently into the same cluster or separating states that are dynamically similar.

4.2 ASSESSMENT OF THE CENTER'S SELECTION

In this section, the effectiveness of the center selection is discussed. The experimental conditions represented by the 17 selected centers are presented in Table 5.

Table 5 – Experimental conditions for the clusters' centers.

Cluster	Center Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content [%]	Author	Data Set
●	10	1310.0	20.0	1.99	$NH_3/N_2/O_2$	0.0	B.Shu	5
●	25	1373.0	20.0	0.5	$NH_3/N_2/O_2$	0.0	B.Shu	1
●	30	1660.8	2.03	1.0	$NH_3/N_2/O_2/H_2/Ar$	50.0	J.Baker_2023	1
●	35	1252.2	1.27	1.0	$NH_3/O_2/H_2/Ar$	70.0	J.Chen_2021	7
●	75	1137.0	9.95	1.0	$NH_3/O_2/H_2/Ar$	70.0	J.Chen_2021	8
●	82	1437.2	1.21	1.0	$NH_3/O_2/H_2/Ar$	30.0	J.Chen_2021	5
●	84	1265.4	9.87	1.0	$NH_3/O_2/H_2/Ar$	30.0	J.Chen_2021	6
●	115	1488.0	14.49	1.0	$NH_3/O_2/Ar$	0.0	Peng_2023	2
●	118	1731.0	12.87	1.0	$NH_3/O_2/Ar$	0.0	Peng_2023	2
●	144	1331.0	21.08	1.0	$NH_3/O_2/Ar$	0.0	Peng_2023	8
●	166	2161.0	1.46	2.0	$NH_3/O_2/Ar$	0.0	O.Mathieu	11
●	175	1969.0	1.59	0.5	$NH_3/O_2/Ar$	0.0	O.Mathieu	2
●	178	2312.0	1.45	0.5	$NH_3/O_2/Ar$	0.0	O.Mathieu	2
●	190	2253.0	1.43	1.0	$NH_3/O_2/Ar$	0.0	O.Mathieu	5
▲	212	1895.0	1.45	1.0	$NH_3/O_2/Ar$	0.0	O.Mathieu	8
▲	236	1742.0	11.15	0.5	$NH_3/O_2/Ar$	0.0	O.Mathieu	3
▲	250	1844.0	28.88	2.0	$NH_3/O_2/Ar$	0.0	O.Mathieu	10

Source: Author (2024).

From a technical perspective, the validation of the algorithm involves verifying its primary goal: identifying centers based on the similarities among reaction networks

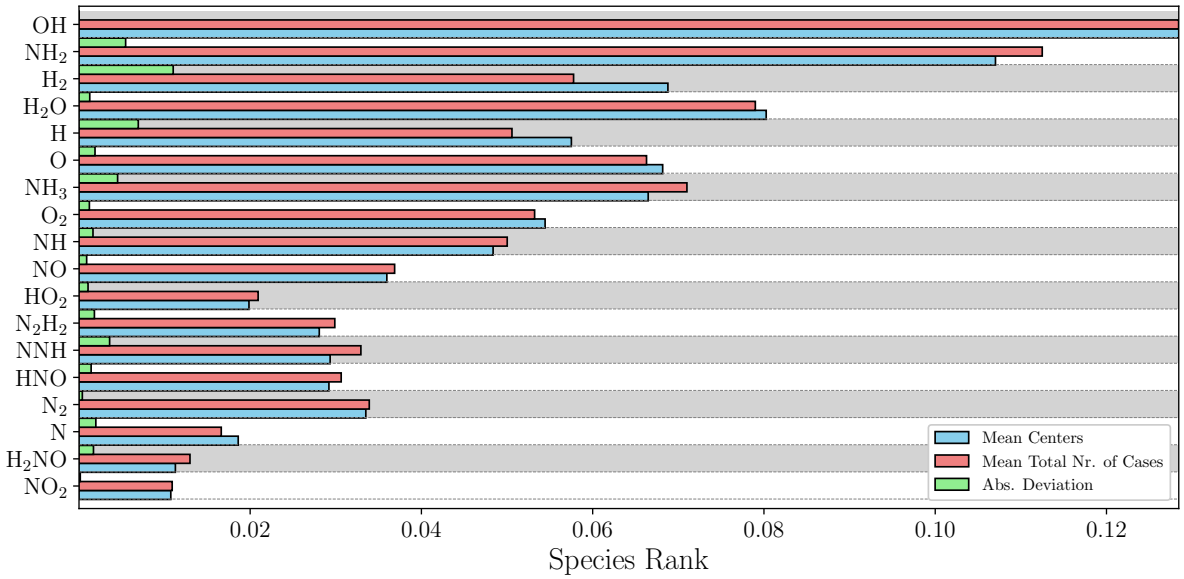
while preserving all relevant information. Intuitively, this objective is achieved if the following condition is satisfied for all species in the chemical kinetic model:

$$\frac{1}{N_c}\sum_{i=1}^{N_c}(SR_{sk}) = \frac{1}{N}\sum_{i=1}^{N}(SR_{sk}) \tag{56}$$

where N_c is the total number of the centers, and N is the total number of cases.

In this perspective, Fig. 9 presents a comparison of the mean values of the species ranked the most important, calculated both across the entire dataset and across the 17 centers.

Figure 9 – Comparison between the mean of the species rank values calculated over the identified centers and all dataset. Results are shown for the most important species.



Source: Author (2024).

In absolute terms, no significative differences occur. However, for some species such as H_2 and H, differences between the mean values are more pronounced.

To investigate the possible cause leading to these discrepancies, the species rank probability distributions across the entire domain of cases are analyzed for all species using the Kernel Density Estimation (KDE) method (PARZEN, 1962). KDE is a non-parametric method to estimate the probability density function of a random variable, by smoothing out the data points (e.g. a histogram) to create a continuous estimate of the probability density. The resulting probability distributions are provided in Appendix A.

It can be observed that species such as H_2 and H, which exhibit high-variance probability distribution, tend to show larger deviations between their mean values. Con-

versely, when the probability distribution presents well-defined primary modes characterized by a smaller variance, the deviations are significantly reduced.

To evaluate the effectiveness of the algorithm in identifying centers corresponding to distinct combustion regimes, the intervals of the properties characterizing the initial conditions covered by a single cluster are shown in Tab. 6.

Table 6 – Initial conditions ranges covered by selected clusters.

Cluster	Φ			Pressure [bar]			Temperature [K]			H_2 Content [%]			Dilution		
	Min	Max	CV	Min	Max	CV	Min	Max	CV	Min	Max	CV	Min	Max	CV
●	1.99	2.00	0.00	16.00	41.90	0.36	1183.00	1561.00	0.10	0.00	0.00	0.00	3.75	10.30	0.55
●	0.50	0.50	0.00	18.95	42.20	0.35	1181.00	1555.00	0.08	0.00	0.00	0.00	3.77	4.68	0.11
●	1.00	1.00	0.00	2.03	2.03	0.00	1511.40	1774.10	0.06	50.00	50.00	0.00	77.40	77.40	0.00
●	1.00	1.00	0.00	1.10	1.28	0.06	1022.70	1600.60	0.18	70.00	70.00	0.00	29.77	29.77	0.00
●	1.00	1.00	0.00	9.35	10.09	0.03	1052.20	1296.90	0.07	70.00	70.00	0.00	29.77	29.77	0.00
●	1.00	1.00	0.00	1.19	12.77	1.06	1199.40	1939.80	0.16	30.00	30.00	0.00	28.74	28.74	0.00
●	1.00	1.00	0.00	9.87	10.81	0.04	1206.60	1451.90	0.07	30.00	30.00	0.00	28.74	28.74	0.00
■	0.50	1.00	0.18	9.45	30.09	0.51	1360.00	1672.00	0.06	0.00	5.00	2.36	11.00	231.04	0.99
■	0.99	1.00	0.00	1.02	29.59	0.80	1499.80	1956.90	0.06	0.00	5.00	1.82	16.71	231.04	0.97
■	1.00	1.00	0.00	10.74	42.20	0.53	1180.00	1581.00	0.07	0.00	0.00	0.00	3.76	11.00	0.44
■	2.00	2.00	0.00	1.37	10.23	1.20	1992.00	2404.00	0.07	0.00	0.00	0.00	363.04	363.04	0.00
■	0.50	0.50	0.00	1.48	11.04	0.96	1923.00	2077.00	0.03	0.00	0.00	0.00	165.00	165.00	0.00
■	0.50	0.50	0.00	1.31	1.48	0.05	2164.00	2479.00	0.05	0.00	0.00	0.00	165.00	165.00	0.00
■	1.00	1.00	0.00	1.27	1.46	0.04	2087.00	2489.00	0.06	0.00	0.00	0.00	114.35	231.04	0.30
▲	0.99	1.00	0.00	1.21	28.27	0.91	1742.00	2085.00	0.05	0.00	5.00	2.82	28.15	231.04	0.51
▲	0.50	0.50	0.00	10.74	29.08	0.50	1626.00	1894.00	0.05	0.00	0.00	0.00	165.00	165.00	0.00
▲	2.00	2.00	0.00	10.64	30.80	0.47	1584.00	1962.00	0.07	0.00	0.00	0.00	363.04	363.04	0.00
Mean CV		0.01			0.46			0.08			0.41			0.23	

Source: Author (2024).

For all the properties, the coefficient of variation (CV) (LOVIE, 2005) within the identified intervals is also presented and its mean value averaged over all 17 clusters is computed. The coefficient of variation is defined as the ratio of the standard deviation to the mean. A smaller CV indicates that the property plays a more significant role in characterizing the underlying chemistry of a center.

In this perspective, the equivalence ratio (ϕ) presents the smallest value. Therefore, a first characterization of the centers is made based on ϕ, which leads to the identification of three major ϕ-regimes corresponding to lean (ϕ = 0.5), stoichiometric (ϕ = 1) and rich (ϕ = 2) conditions.

At parity of ϕ, different regimes can be identified based on different temperature ranges and dilution levels.

- ϕ = 0.5
  - ●25: $T = 1181 - 1555$ K;
  - ▲236: $T = 1626 - 1894$ K;
  - ■175: $T = 1923 - 2077$ K;
  - ■178: $T = 2164 - 2479$ K;
- ϕ = 1.0

  - ■144: $T = 1180 - 1581$ K; Low Dilution;
  - ■115: $T = 1360 - 1672$ K; High Dilution;
  - ■118: $T = 1499 - 1956$ K;
  - ▲212: $T = 1742 - 2085$ K;
  - ■190: $T = 2087 - 2489$ K;

- ϕ = 2.0
  - ●10: $T = 1183 - 1561$ K;
  - ▲250: $T = 1584 - 1962$ K;
  - ■166: $T = 1992 - 2404$ K;

The temperature overlap observed for centers 115, 118, and 212 arises from cases characterized by an initial 5% H_2 content, which have been grouped with cases involving pure NH_3. Consistent with the findings of Chen et al. (CHEN et al., 2021), a 5% H_2 content within the initial composition, does not alter the overall reaction flux compared to pure NH_3 cases.

However, the presence of H_2 in the mixture enhances the overall reactivity, shifting it towards lower temperatures. This shift explains the temperature interval overlap observed between the three centers.

In a low-to-intermediate temperature regime ($T = 1100 - 1500$ K), the reaction H + O_2 + M = HO_2 + M is an important source of HO_2 radicals, which are majorly responsible in influencing reactivity under this regime. At moderately low pressures, the reaction rate is approximately proportional to diluent concentrations. The collision efficiency, which is the proportionality factor, determines the production of HO_2. Adjusting dilution levels or the composition of diluents affects reactivity, as shown by distinct clusters corresponding to low (144) and high dilution levels (115).

For these cases, because of the relatively high CV value of the pressure, it can be concluded that its influence on variations in the underlying chemistry is negligible.

The remaining centers are classified based on the initial H_2-content and initial pressure values.

- H_2 content = 30%
  - ●82: $T = 1199 - 1825$ K; $P = 1$ bar; $T = 1551 - 1939$ K; $P = 10$ bar;
  - ●84: $T = 1206 - 1451$ K; $P = 10$ bar;
- H_2 content = 50%
  - ●30: $T = 1511 - 1774$ K; ϕ = 2; $P = 2$ bar
- H_2 content = 70%
  - ●75: $T = 1052 - 1296$ K; $P = 10$ bar
  - ●35: $T = 1022 - 1600$ K; $P = 1$ bar

Considering centers 82 and 84, which represent cases with an initial 30% H_2 content, it can be concluded that pressure plays a significant role in the low-to-intermediate temperature range. Specifically, for center 82 at $P = 1$ bar, low-temperature data appears to cluster together with high-temperature data. In contrast, at $P = 10$ bar, low-temperature data forms a distinct cluster.

This behavior can be partially explained by findings from Chen et al. (CHEN et al., 2021), which indicate that at 30% H_2 content, the binary mixture's shift toward hydrogen chemistry causes the ignition delay time to exhibit highly nonlinear behavior with respect to temperature. This effect becomes increasingly pronounced at higher pressures. Nonetheless, further investigation is necessary to fully understand the chemistry leading to the differentiation between these clusters.

All the cases grouped in the presented centers are reported in Appendix B.

To determine which species exhibit rank variations that significantly influence the center selection, an $Nc \times Ns$ dissimilarity matrix ($\mathbf{D}$) is introduced, whose elements (d_{i,s_k}) are given by:

$$d_{i,s_k} = \frac{1}{(N_C - 1)} \sum_{j=1, j \neq i}^{N_c} \text{JSD}(pr_{i,s_k} || pr_{j,s_k}) \tag{57}$$

The obtained matrix is depicted in Fig. 10. The columns are sorted in descending order based on:

$$d_{i,s_k} = \frac{1}{(N_C)} \sum_{i=1}^{N_C} d_{i,s_k} \tag{58}$$

The clustering algorithm primarily distinguishes cases by prioritizing differences in the highest-ranking species. Notably, the largest deviations are observed for the H_2 species. This aligns with the fact that the selected centers correspond to conditions with varying initial H_2 content.

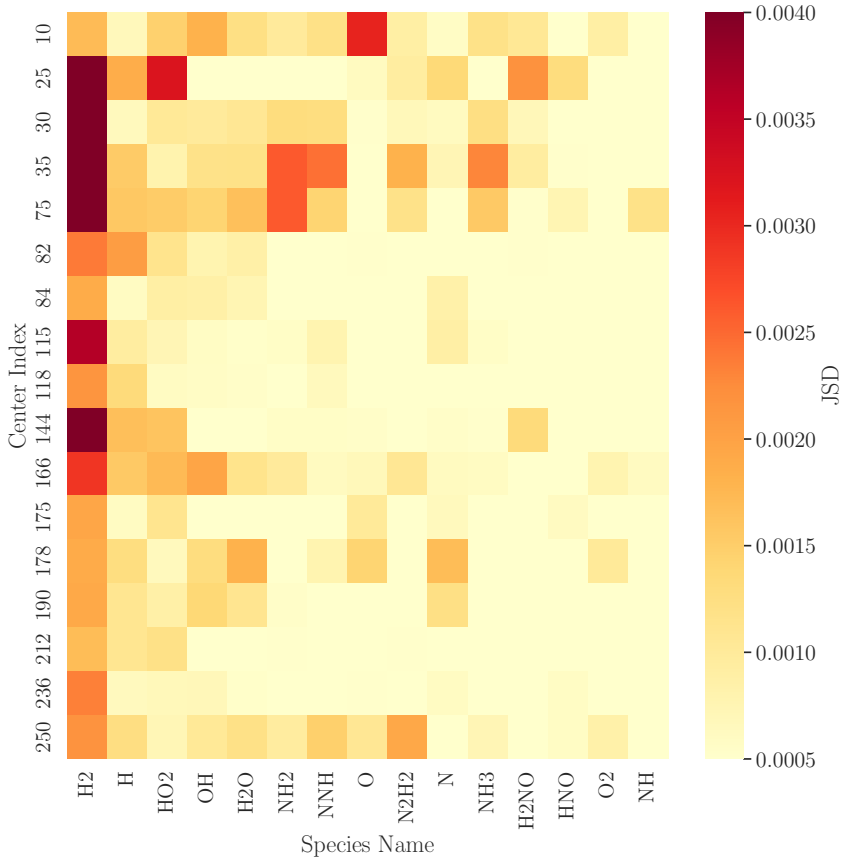
As the initial H_2 concentration increases, H_2 chemistry becomes more influential within the reaction network, resulting in greater variations in the H_2 species rank. Additional deviations are associated with reactivity-controlling radicals such as H, OH, and HO_2, or species such as N_2H_2 and HNO whose chemistry was found to be particularly influenced by changes in the initial ϕ value.

Through the proposed method, it is possible to identify the combustion regimes encompassed by a broad database of experimental conditions, thus enabling a comprehensive analysis of the chemistry governing the combustion process.

4.3 APPLICATION TO SKELETAL REDUCTION

The clustering is then applied to conduct the reduction process by only considering centers as representative cases of the detailed model validity range. In this

Figure 10 – Dissimilarity matrix depicting the average value over the centers of the deviation between components of species rank vectors.



Source: Author (2024).

perspective, the final rank associated with the species is given by Eq. 49.

Species are then sorted in ascending order based on their ranking and removed sequentially. The threshold on the minimum number of species (N_s, min) is set to 21. For all the generated reduced models, the curve matching values between the results obtained with the reduced model obtained with the 17 clusters and the detailed model is computed and presented in Tab. 7. The first row presents a comparison between the detailed model and experimental data. The color scheme is used to represent the quality of the reduced model, with green indicating good agreement between the reduced model and detailed model, while yellow and red indicate poorer agreement.

The study from Girhe et al. (2024) shows that the NUIG_2023 model agrees well with the IDT data, particularly at low fuel concentrations. However, none of the models examined can consistently predict the IDT measurements across the full range of fuel concentrations. While there are many kinetic models available, their validations are typically performed under different and limited conditions. This variability in conditions helps explain the discrepancies observed when comparing the detailed model

with experimental data from different authors.

Table 7 – Curve Matching Score: Comparison of the reduced model (based on the clusters' centers) with the detailed model.

Nr. Species	Removed Species	B.Shu	Peng 2023	J.Chen 2021	O.Mathieu	J.Baker 2023	Mean CM
43	/	0.935455	0.910016	0.925564	0.949141	0.893846	0.929755
38	CO CO2 CH4 C2H6 HE	0.9954	0.9942	0.9936	0.9936	0.9973	0.9942
37	ONHN	0.9957	0.9942	0.9936	0.9934	0.9967	0.9941
36	T-HNN(O)OH	0.9953	0.9942	0.9936	0.9936	0.9977	0.9942
35	HONO2	0.9953	0.9942	0.9936	0.9937	0.9971	0.9942
34	HNO2	0.9951	0.9941	0.9936	0.9936	0.9972	0.9941
33	HON	0.9949	0.9941	0.9936	0.9936	0.9977	0.9941
32	C-HNN(O)OH	0.9954	0.9941	0.9936	0.9936	0.9971	0.9942
31	C-ONNH	0.9955	0.9940	0.9936	0.9936	0.9955	0.9941
30	NO3	0.9956	0.9940	0.9936	0.9937	0.9962	0.9941
29	T-ONNH	0.9946	0.9935	0.9936	0.9936	0.9967	0.9938
28	H2NNO2	0.9945	0.9934	0.9936	0.9936	0.9964	0.9938
27	NH2OH	0.9938	0.9919	0.9932	0.9934	0.9964	0.9931
26	HONO	0.9873	0.9858	0.9929	0.9933	0.9982	0.9903
25	H2NN	0.9891	0.9864	0.9930	0.9926	0.9957	0.9905
24	N2H4	0.9879	0.9850	0.9890	0.9926	0.9943	0.9891
23	N2H3	0.9871	0.9833	0.9884	0.9928	0.9945	0.9884
22	H2O2	0.9461	0.9404	0.9736	0.9907	0.9942	0.9663
21	HNOH	0.9486	0.9415	0.9740	0.9907	0.9948	0.9671

Source: Author (2024).

To validate the robustness of the proposed method for skeletal reduction applications, a performance comparison of the reduced models obtained by considering all cases is also presented in Table 8.

Table 8 – Curve Matching Score: Comparison of the reduced model (based on all cases) with the detailed model.

Nr. Species	Removed Species	B.Shu	Peng 2023	J.Chen 2021	O.Mathieu	J.Baker 2023	Mean CM
43	/	0.935455	0.910016	0.925564	0.949141	0.893846	0.929755
38	CO CO2 CH4 C2H6 HE	0.9954	0.9942	0.9936	0.9936	0.9973	0.9942
37	ONHN	0.9957	0.9942	0.9936	0.9934	0.9967	0.9941
36	HONO2	0.9953	0.9942	0.9936	0.9936	0.9973	0.9941
35	T-HNN(O)OH	0.9953	0.9942	0.9936	0.9937	0.9971	0.9942
34	HON	0.9952	0.9942	0.9936	0.9937	0.9973	0.9942
33	HNO2	0.9949	0.9941	0.9936	0.9936	0.9977	0.9941
32	C-HNN(O)OH	0.9954	0.9941	0.9936	0.9936	0.9971	0.9942
31	C-ONNH	0.9955	0.9940	0.9936	0.9936	0.9955	0.9941
30	NO3	0.9956	0.9940	0.9936	0.9937	0.9962	0.9941
29	NH2OH	0.9944	0.9926	0.9932	0.9934	0.9974	0.9934
28	H2NNO2	0.9944	0.9926	0.9932	0.9934	0.9957	0.9933
27	T-ONNH	0.9938	0.9919	0.9932	0.9934	0.9964	0.9931
26	HONO	0.9873	0.9858	0.9929	0.9933	0.9982	0.9903
25	H2NN	0.9891	0.9864	0.9930	0.9926	0.9957	0.9905
24	N2H4	0.9879	0.9850	0.9890	0.9926	0.9943	0.9891
23	N2H3	0.9871	0.9833	0.9884	0.9928	0.9945	0.9884
22	HNOH	0.9835	0.9801	0.9884	0.9928	0.9943	0.9869
21	H2O2	0.9486	0.9415	0.9740	0.9907	0.9948	0.9671

Source: Author (2024).

When considering all cases, the species H_2O_2 ranks higher than HNOH, which leads to greater accuracy in the reduced mechanism compared to the case where only

the centers are considered. However, the higher ranking assigned to H_2O_2 comes from the O. Mathieu datasets (MATHIEU; PETERSEN, 2015), which, as seen in the table, appear to be less influenced by the removal H_2O_2.

In general, upon examining the probability distributions provided in the appendix, HNOH presents a higher ranking than H_2O_2. As discussed in the previous section, the clustering algorithm tends to focus on deviations among the highest-ranking species, making it understandable that information on H_2O_2's ranking may be overlooked. This is because a ranking based on reaction flux does not account for the direct effects of removing a species—and the associated reactions—on the quantity of interest (in this case, the ignition delay time). Consequently, such situations are more likely to occur. Possible solutions to address these issues are presented in the next section.

5 CONCLUSION

In this thesis, a method for the automated identification of simulation targets for chemical kinetic mechanism reduction was developed, aiming at identifying similarities between different experimental conditions based on reaction network comparison. A database of 264 auto-ignition experiments for ammonia and ammonia/hydrogen blends was selected as an application case for the proposed method.

The study started with the combustion kinetics simulations, computation of the direct interaction coefficients, and the application of the Directed Relation Graph-based SpeciesRank (DRGSR) method.

Once the similarities between reaction networks were quantified, the affinity propagation algorithm was employed and 17 cluster centers were selected as representatives of the entire database. It was demonstrated that the clustering approach detects similarities based on variations in the rank of the highly-ranked species.

To assess the effectiveness of the proposed method in practical reduction applications, the centers were employed as simulation targets for the reduction process. Overall, the obtained reduced mechanisms showed good agreement compared to detailed models, as quantified by the curve matching scores.

However, a comparison with the performance of the reduced models obtained by considering all the cases in the validation database highlighted some limitations of the proposed method when applied to the reduction process. These discrepancies can be attributed to the fact that the algorithm tends to form clusters based on the overall highly ranked species. As a result, variations in the chemistry of species that are generally less influential may go undetected.

Although the reduction process focused on ignition delay time as the primary validation target, the generalization of the algorithm needs to be further explored by evaluating other combustion properties. Therefore, the following recommendations are proposed for future work:

- Extending the reduction methodology to include other combustion properties, such as laminar flame speed.
- Investigating the impact of different preference selections in the affinity propagation algorithm in clustering results.
- Developing a weighting factor for species, accounting for cases where species with low importance indices may participate in reactions with high sensitivity to the objective function, such as IDT.

BIBLIOGRAPHY

BAKER, J. B. et al. Experimental ignition delay time measurements and chemical kinetics modeling of hydrogen/ammonia/natural gas fuels. **Journal of Engineering for Gas Turbines and Power**, v. 145, n. 4, p. 041002, 12 2022. ISSN 0742-4795. Available on: https://doi.org/10.1115/1.4055721.

BERNARDI, M. et al. Curve matching, a generalized framework for models/experiments comparison: An application to n-heptane combustion kinetic mechanisms. **Combustion and Flame**, v. 168, p. 186–203, 2016. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218016300360.

BLAIN, L. **First ammonia-powered jet flight in 2023: A roadmap to clean aviation**. 2022. New Atlas. Accessed: 2024-12-06. Available on: https://newatlas.com/aircraft/aviation-h2-ammonia-fuel-jet-aircraft/.

BRIN, S.; PAGE, L. The anatomy of a large-scale hypertextual web search engine. **Computer Networks**, v. 30, p. 107–117, 1998. Available on: http://www-db.stanford.edu/~backrub/google.html.

Cambridge Dictionary. **Combustion**. 2024. Available on: https://dictionary.cambridge.org/dictionary/english/combustion. Acesso em: 12 nov. 2024.

CHEN, J.; JIANG, X.; QIN, X.; HUANG, Z. Effect of hydrogen blending on the high temperature auto-ignition of ammonia at elevated pressure. **Fuel**, v. 287, p. 119563, 2021. ISSN 0016-2361. Available on: https://www.sciencedirect.com/science/article/pii/S001623612032559X.

CHOUDHARY, R. et al. Shock tube study of ethanol pyrolysis ii: Rate constant measurements and modeling. **Combustion and Flame**, v. 233, p. 111554, 2021. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218021002972.

COHEN, S. D.; HINDMARSH, A. C.; DUBOIS, P. F. CVODE, a stiff/nonstiff ODE solver in C. **Computers in Physics**, v. 10, n. 2, p. 138–143, 1996. Available on: https://aip.scitation.org/doi/abs/10.1063/1.4822377.

EFRON, B.; TIBSHIRANI, R. J. **An Introduction to the Bootstrap**. New York: Chapman & Hall/CRC, 1994. ISBN 978-0412042317.

ERDEMIR, D.; DINCER, I. A perspective on the use of ammonia as a clean fuel: Challenges and solutions. **International Journal of Energy Research**, v. 45, n. 4, p. 4827–4834, 2021. Available on: https://onlinelibrary.wiley.com/doi/abs/10.1002/er.6232.

FREY, B. J.; DUECK, D. Clustering by passing messages between data points. **Science**, v. 315, n. 5814, p. 972–976, 2007. Available on: https://www.science.org/doi/abs/10.1126/science.1136800.

GARCIA, A. C.; SANTOS, R. G. dos. Reduction of chemical kinetics mechanisms: a review of methods and applications in the development of new fuels. *In*: **SAE BRASIL 2021 Web Forum**. SAE International, 2022. Available on: https://doi.org/10.4271/2021-36-0031.

GIDDEY, S.; BADWAL, S.; KULKARNI, A. Review of electrochemical ammonia production technologies and materials. **International Journal of Hydrogen Energy**, v. 38, n. 34, p. 14576–14594, 2013. ISSN 0360-3199. Available on: https://www.sciencedirect.com/science/article/pii/S0360319913022775.

GIRHE, S. et al. Ammonia and ammonia/hydrogen combustion: Comprehensive quantitative assessment of kinetic models and examination of critical parameters. **Combustion and Flame**, v. 267, p. 113560, 2024. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218024002694.

GOKULAKRISHNAN, P.; LAWRENCE, A.; MCLELLAN, P.; GRANDMAISON, E. A functional-pca approach for analyzing and reducing complex chemical mechanisms. **Computers Chemical Engineering**, v. 30, n. 6, p. 1093–1101, 2006. ISSN 0098-1354. Available on: https://www.sciencedirect.com/science/article/pii/S0098135406000299.

KHALED, F.; BADRA, J.; FAROOQ, A. Ignition delay time correlation of fuel blends based on livengood-wu description. **Fuel**, v. 209, p. 776–786, 2017. ISSN 0016-2361. Available on: https://www.sciencedirect.com/science/article/pii/S0016236117309572.

KULLBACK, S.; LEIBLER, R. A. On Information and Sufficiency. **The Annals of Mathematical Statistics**, Institute of Mathematical Statistics, v. 22, n. 1, p. 79 – 86, 1951. Available on: https://doi.org/10.1214/aoms/1177729694.

LACKNER, M.; PALOTÁS, Á.; WINTER, F. **Combustion: From Basics to Applications**. Wiley, 2013. ISBN 9783527333769. Available on: https://books.google.de/books?id=k632JE0fLSwC.

LEE, D. et al. Analysis of start-up behavior based on the dynamic simulation of an sofc–engine hybrid system. **Energy Conversion and Management**, v. 272, p. 116384, 2022. ISSN 0196-8904. Available on: https://www.sciencedirect.com/science/article/pii/S0196890422011621.

LOVIE, P. Coefficient of variation. *In*: ____. **Encyclopedia of Statistics in Behavioral Science**. John Wiley Sons, Ltd, 2005. ISBN 9780470013199. Available on: https://onlinelibrary.wiley.com/doi/abs/10.1002/0470013192.bsa107.

LU, T.; LAW, C. K. On the applicability of directed relation graphs to the reduction of reaction mechanisms. **Combustion and Flame**, v. 146, n. 3, p. 472–483, 2006. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218006001398.

MATHIEU, O.; PETERSEN, E. L. Experimental and modeling study on the high-temperature oxidation of ammonia and related nox chemistry. **Combustion and Flame**, v. 162, n. 3, p. 554–570, 2015. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218014002752.

MCBRIDE, B. **NASA Glenn Coefficients for Calculating Thermodynamic Properties of Individual Species**. National Aeronautics and Space Administration, John H. Glenn Research Center at Lewis Field, 2002. (NASA Technical Paper). Available on: https://books.google.de/books?id=TAEVAQAAIAAJ.

NASA. **X-15 Hypersonic Research Program**. 2024. Accessed: 2024-12-06. Available on: https://www.nasa.gov/reference/x-15/.

NIELSEN, F. On the jensen–shannon symmetrization of distances relying on abstract means. **Entropy**, v. 21, n. 5, 2019. ISSN 1099-4300. Available on: https://www.mdpi.com/1099-4300/21/5/485.

NISHIDA, M. Chapter 4.1 - shock tubes. *In*: BEN-DOR, G.; IGRA, O.; ELPERIN, T. (Ed.). **Handbook of Shock Waves**. Burlington: Academic Press, 2001. p. 553–585. ISBN 978-0-12-086430-0. Available on: https://www.sciencedirect.com/science/article/pii/B9780120864300500129.

PAGE, L.; BRIN, S.; MOTWANI, R.; WINOGRAD, T. **The PageRank Citation Ranking: Bringing Order to the Web**. [S.l.], 1998. Available on: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.31.1768.

PARZEN, E. On Estimation of a Probability Density Function and Mode. **The Annals of Mathematical Statistics**, Institute of Mathematical Statistics, v. 33, n. 3, p. 1065 – 1076, 1962. Available on: https://doi.org/10.1214/aoms/1177704472.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PENG, Y.; RANJAN, D.; SUN, W. A shock tube study of fuel concentration effect on high-pressure autoignition delay of ammonia. **Applications in Energy and Combustion Science**, v. 16, p. 100202, 2023. ISSN 2666-352X. Available on: https://www.sciencedirect.com/science/article/pii/S2666352X23000912.

PEPIOT-DESJARDINS, P.; PITSCH, H. An efficient error-propagation-based reduction method for large chemical kinetic mechanisms. **Combustion and Flame**, v. 154, n. 1, p. 67–81, 2008. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218007003264.

PITSCH, H. **FlameMaster**: A c++ computer program for 0d combustion and 1d laminar flame calculations. 1993. Available on: https://www.itv.rwth-aachen.de/downloads/flamemaster/.

RAMALLI, E. et al. Automatic validation and analysis of predictive models by means of big data and data science. **Chemical Engineering Journal**, v. 454, p. 140149, 2023. ISSN 1385-8947. Available on: https://www.sciencedirect.com/science/article/pii/S1385894722056297.

GANGOLI RAO, A. et al. **The redesign of an A320 using ammonia as the energy carrier**. 2022. Available on: http://dx.doi.org/10.13140/RG.2.2.19143.06561.

REACTION DESIGN. **Chemkin Theory Manual 17.8**. San Diego, 2016.

RITCHIE, H. What share of global co emissions come from aviation? **Our World in Data**, 2024. Https://ourworldindata.org/global-aviation-emissions.

RITCHIE, H.; ROSADO, P.; ROSER, M. Breakdown of carbon dioxide, methane and nitrous oxide emissions by sector. **Our World in Data**, 2020. Available on: https://ourworldindata.org/emissions-by-sector.

SHU, B. et al. A shock tube and modeling study on the autoignition properties of ammonia at intermediate temperatures. **Proceedings of the Combustion Institute**, v. 37, n. 1, p. 205–211, 2019. ISSN 1540-7489. Available on: https://www.sciencedirect.com/science/article/pii/S1540748918304929.

SUN, W.; CHEN, Z.; GOU, X.; JU, Y. A path flux analysis method for the reduction of detailed chemical kinetic mechanisms. **Combustion and Flame**, v. 157, n. 7, p. 1298–1307, 2010. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218010000842.

SUNG, C.-J.; CURRAN, H. J. Using rapid compression machines for chemical kinetics studies. **Progress in Energy and Combustion Science**, v. 44, p. 1–18, 2014. ISSN 0360-1285. Available on: https://www.sciencedirect.com/science/article/pii/S0360128514000148.

THALER, B. **The Mathematics of Google's PageRank**. 2013. https://www.ams.org/publicoutreach/feature-column/fcarc-pagerank. Accessed: June 10, 2024.

TURNS, S. **An Introduction to Combustion: Concepts and Applications**. McGraw-Hill, 2000. (McGraw-Hill series in mechanical engineering). ISBN 9780071169103. Available on: https://books.google.de/books?id=rzo8PgAACAAJ.

United Nations. **Causes and Effects of Climate Change**. 2024. Accessed: 2024-02-08. Available on: https://www.un.org/en/climatechange/science/causes-effects-climate-change#:~:text=Fossil%20fuels%20%E2%80%93%20coal%2C%20oil%20and,of%20all%20carbon%20dioxide%20emissions.

VIVEK, P.; SITHARAM, T. G. Fundamental of shock waves and shock tube. *In*: ____. **Granular Materials Under Shock and Blast Loading**. Singapore: Springer Singapore, 2020. p. 11–19. ISBN 978-981-15-0438-9.

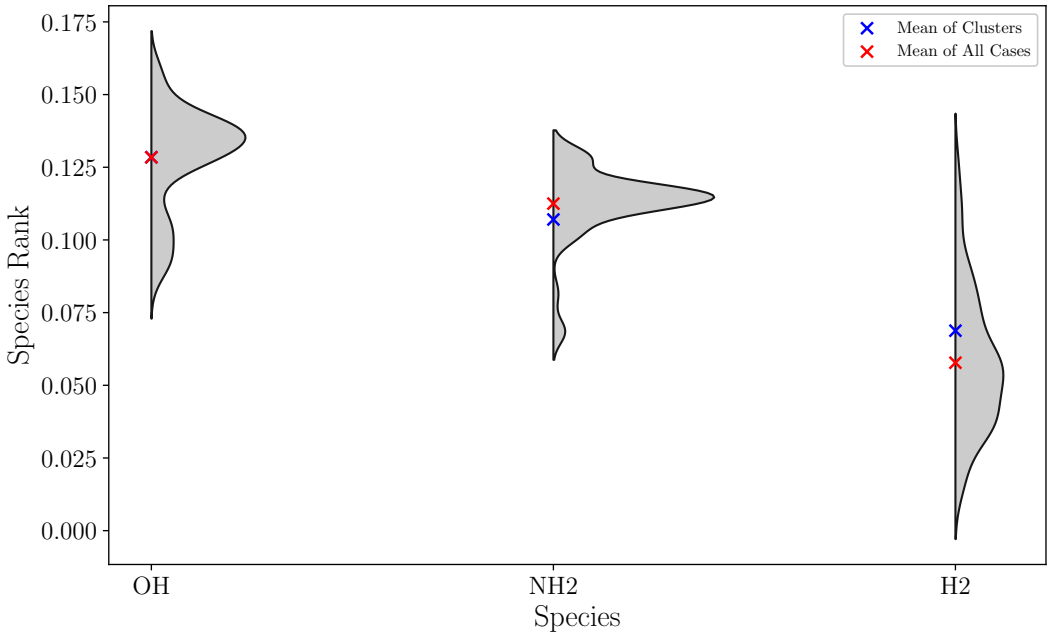
WANG, Y.; WU, Y.; LIN, S.; LAW, C. K.; YANG, B. An efficient mechanism reduction algorithm: Directed Relation Graph-based SpeciesRank (DRGSR). **Combustion and Flame**, Submitted, 2024.

WARNATZ, J.; MAAS, U.; DIBBLE, R. **Combustion: Physical and Chemical Fundamentals, Modelling and Simulation, Experiments, Pollutant Formation.** Springer, 2001. ISBN 9783540677512. Available on: https://books.google.de/books?id=hldApAvyZf4C.

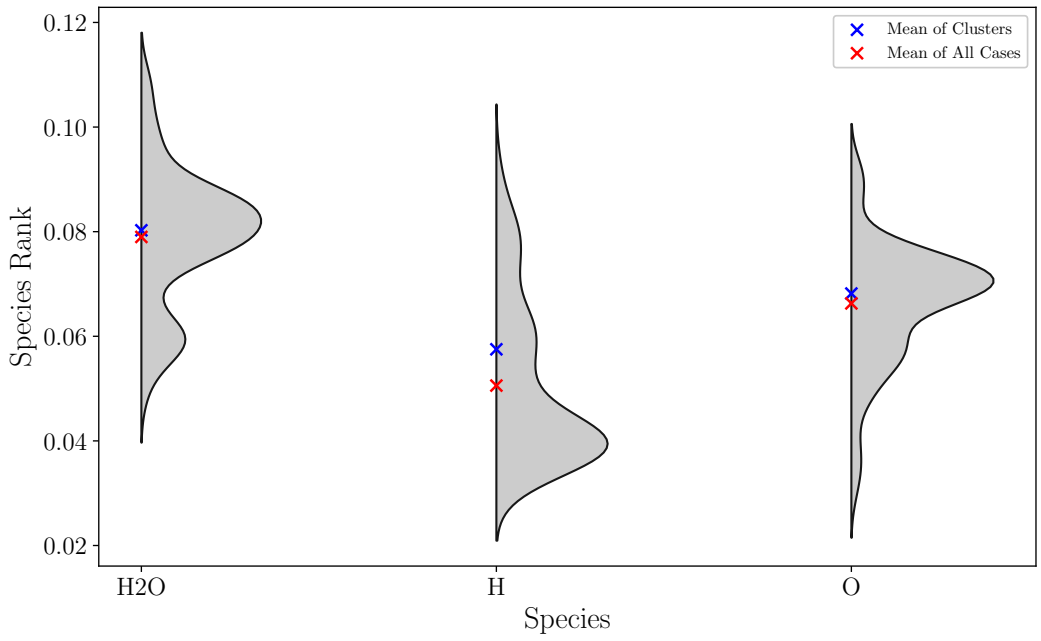
ZHU, Y. et al. The combustion chemistry of ammonia and ammonia/hydrogen mixtures: A comprehensive chemical kinetic modeling study. **Combustion and Flame**, v. 260, p. 113239, 2024. ISSN 0010-2180. Available on: https://www.sciencedirect.com/science/article/pii/S0010218023006132.

Appendix A – DISTRIBUTION OF SPECIES RANKING

Figure 11 – Distribution of ranking scores for each species across all cases – Part 1.



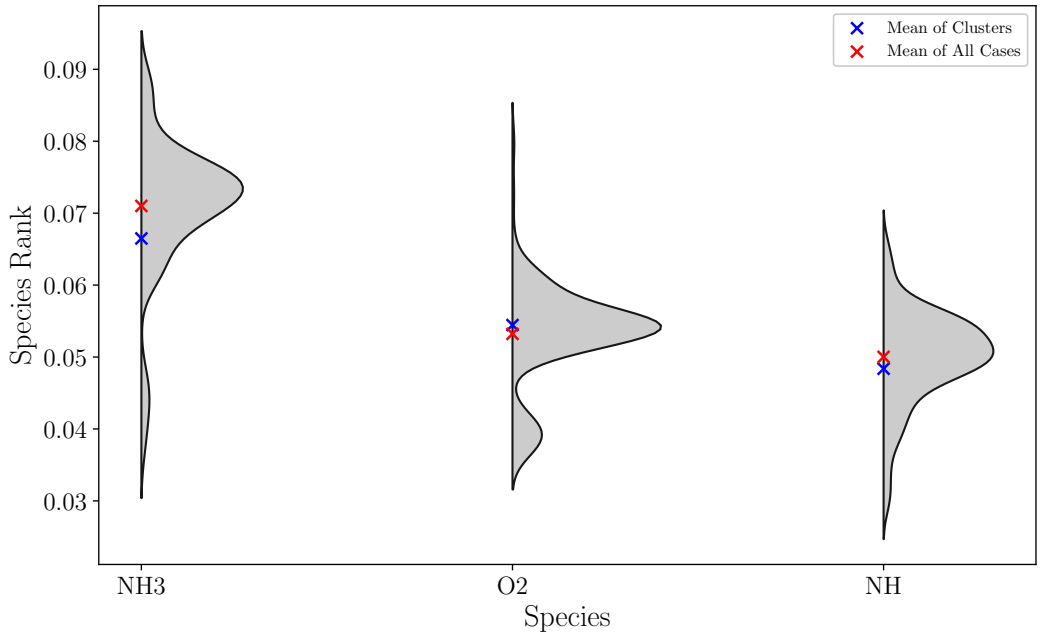
(a)



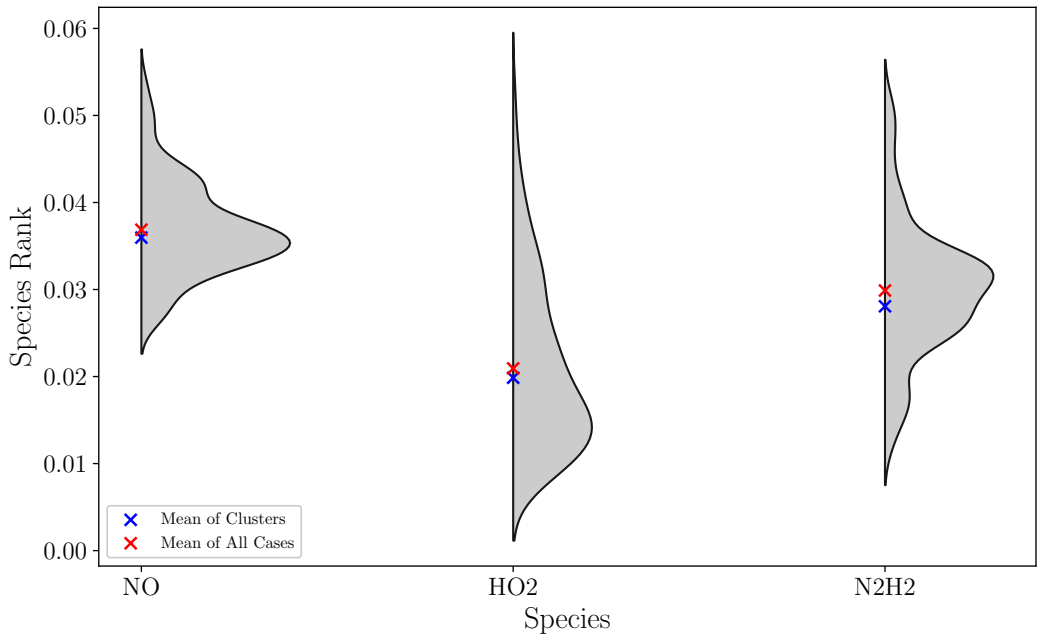
(b)

Source: Author (2024).

Figure 12 – Distribution of ranking scores for each species across all cases – Part 2.



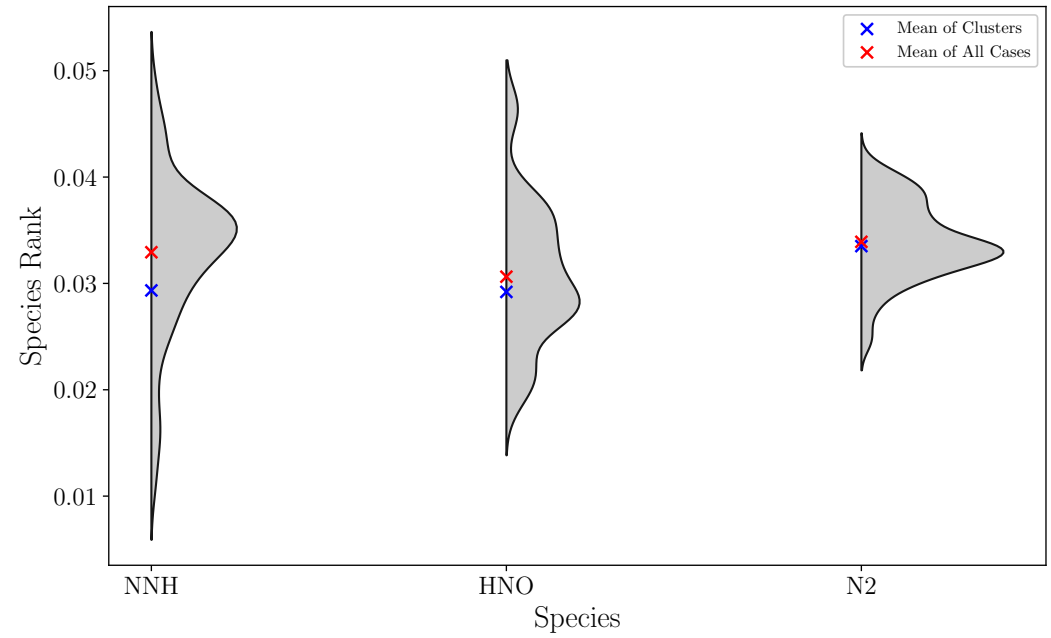
(a)



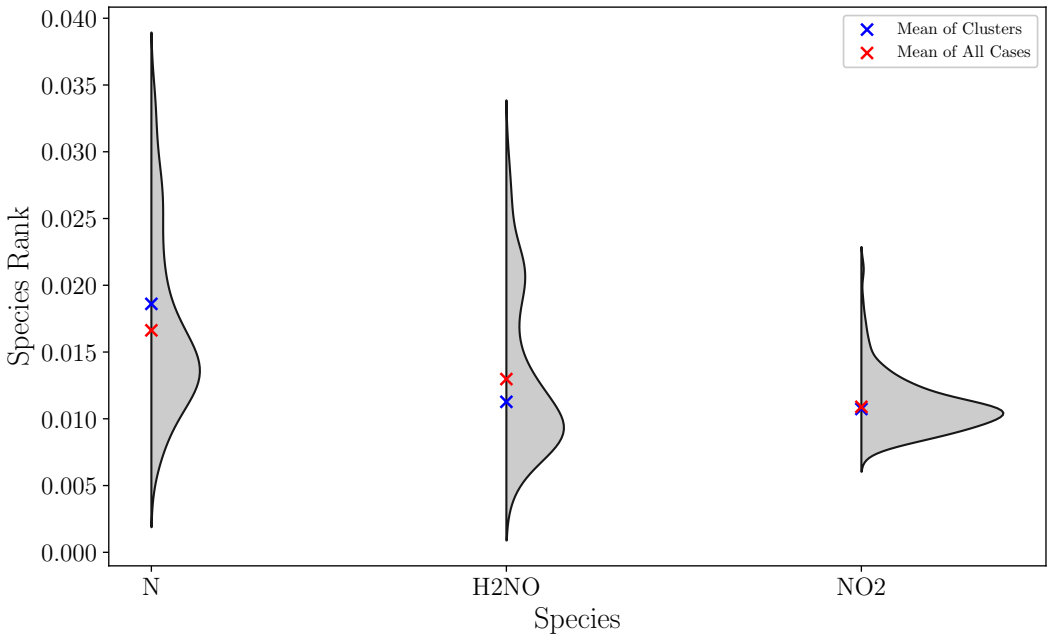
(b)

Source: Author (2024).

Figure 13 – Distribution of ranking scores for each species across all cases – Part 3.



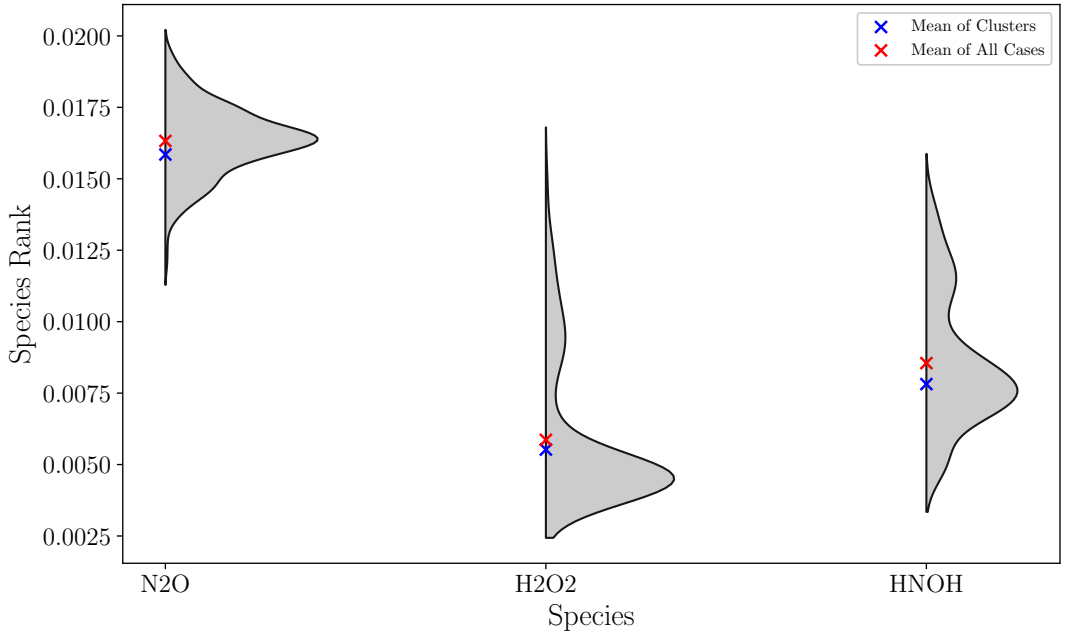
(a)



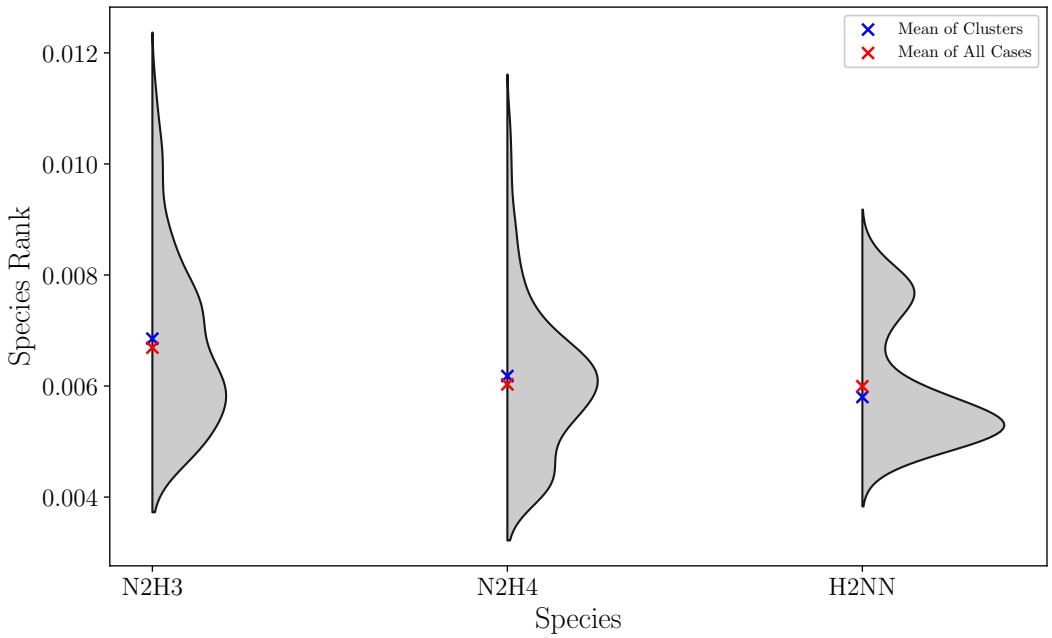
(b)

Source: Author (2024).

Figure 14 – Distribution of ranking scores for each species across all cases – Part 4.



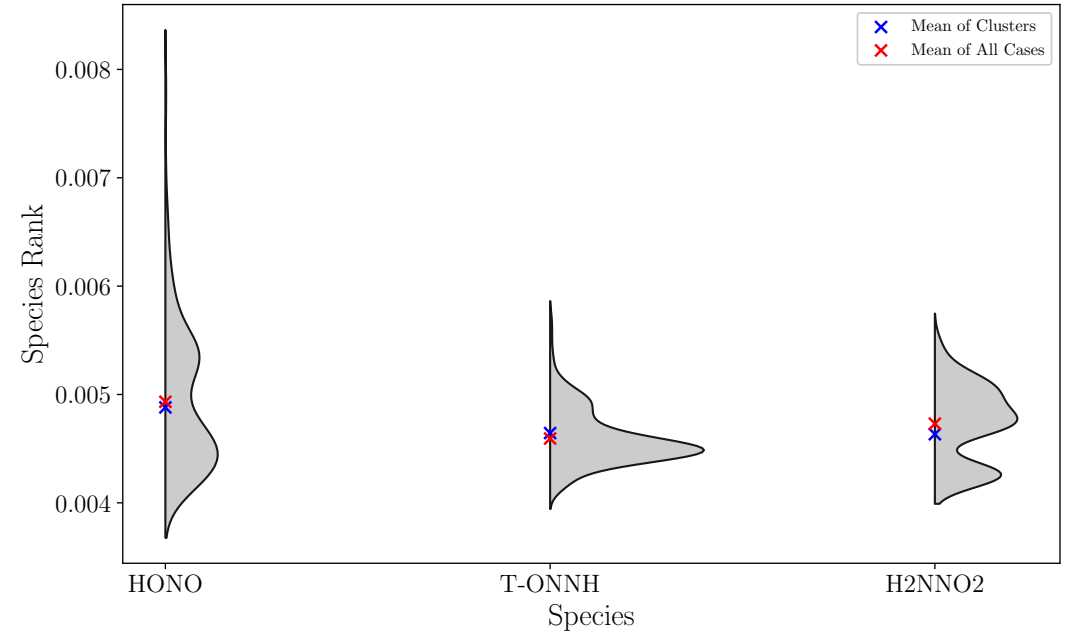
(a)



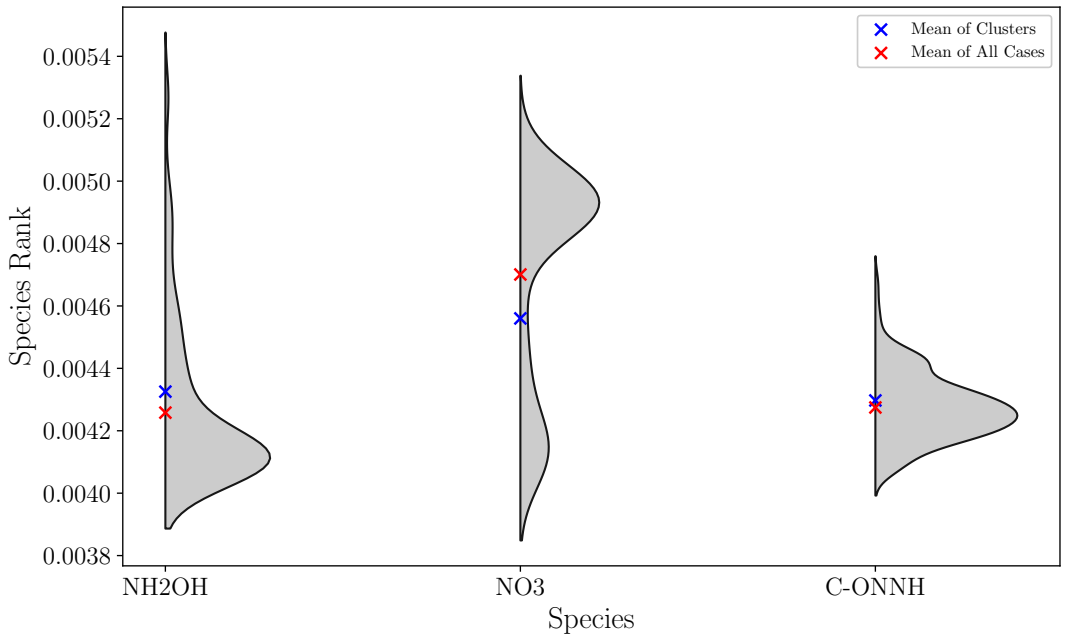
(b)

Source: Author (2024).

Figure 15 – Distribution of ranking scores for each species across all cases – Part 5.



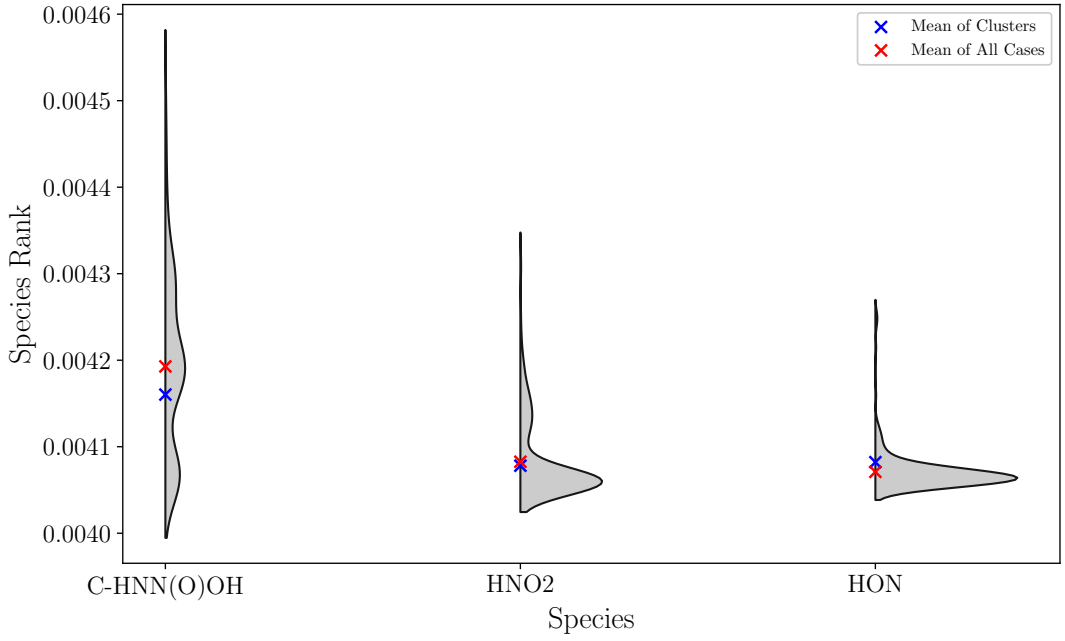
(a)



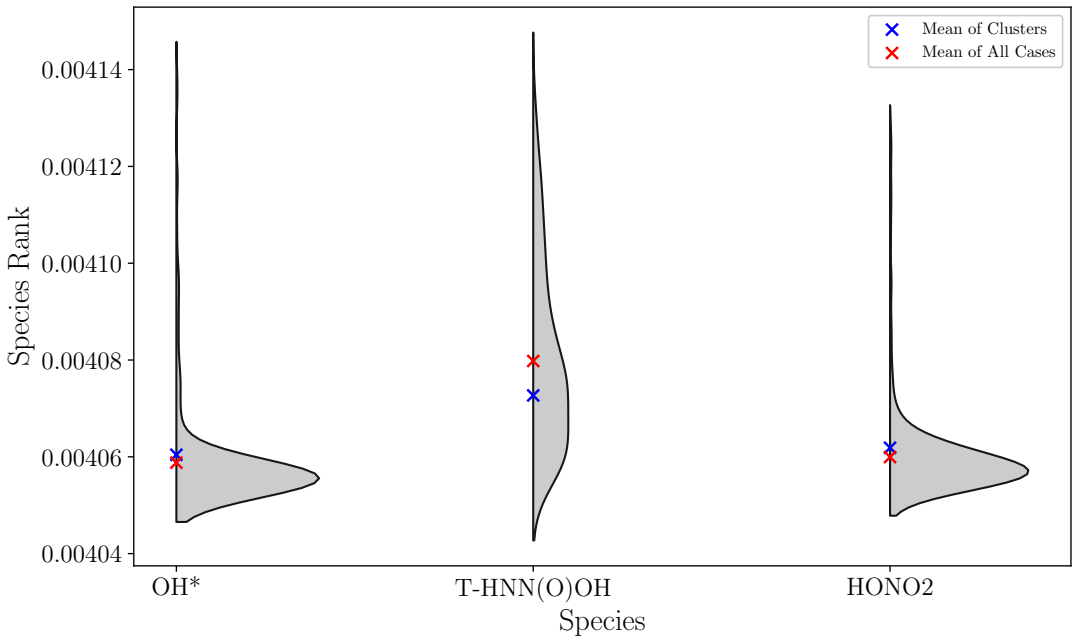
(b)

Source: Author (2024).

Figure 16 – Distribution of ranking scores for each species across all cases – Part 6.



(a)



(b)

Source: Author (2024).

Appendix B – EXPERIMENTAL CONDITIONS FOR ALL DATASET

For each of the following tables 9–25, the first entry, identified by the cluster symbol in the first column, corresponds to the cluster center.

Table 9 – Experimental conditions for cluster 0.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
●	10	1310.0	20.0	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	5
	11	1543.0	19.0	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	5
	12	1290.0	21.0	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	5
	13	1433.0	20.0	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	5
	14	1183.0	16.0	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	5
	15	1301.0	38.8	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	6
	16	1554.0	38.6	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	6
	17	1243.0	40.4	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	6
	18	1416.0	37.6	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	6
	19	1195.0	41.9	1.99	$NH_3/N_2/O_2$	0.0	3.75	B.Shu	6
	152	1352.0	21.99	2.0	$NH_3/O_2/Ar$	0.0	10.3	Peng_2023	9
	153	1405.0	21.89	2.0	$NH_3/O_2/Ar$	0.0	10.3	Peng_2023	9
	154	1561.0	21.18	2.0	$NH_3/O_2/Ar$	0.0	10.3	Peng_2023	9

Source: Author (2024).

Table 10 – Experimental conditions for cluster 1.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
●	25	1373.0	20.0	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	1
	0	1362.0	42.2	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	2
	1	1524.0	38.2	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	2
	2	1272.0	41.2	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	2
	3	1430.0	40.3	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	2
	4	1181.0	38.0	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	2
	26	1492.0	19.7	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	1
	27	1295.0	19.3	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	1
	28	1441.0	20.0	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	1
	29	1213.0	20.0	0.5	$NH_3/N_2/O_2$	0.0	3.77	B.Shu	1
	101	1374.0	22.29	0.5	$NH_3/O_2/Ar$	0.0	4.68	Peng_2023	7
	102	1302.0	22.39	0.5	$NH_3/O_2/Ar$	0.0	4.68	Peng_2023	7
	103	1459.0	21.68	0.5	$NH_3/O_2/Ar$	0.0	4.68	Peng_2023	7
	104	1335.0	22.29	0.5	$NH_3/O_2/Ar$	0.0	4.68	Peng_2023	7
	105	1555.0	18.95	0.5	$NH_3/O_2/Ar$	0.0	4.68	Peng_2023	7

Source: Author (2024).

Table 11 – Experimental conditions for cluster 2.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
●	30	1660.8	2.03	1.0	$NH_3/N_2/O_2/H_2/Ar$	0.5	77.4	J.Baker_2023	1
	31	1511.4	2.03	1.0	$NH_3/N_2/O_2/H_2/Ar$	0.5	77.4	J.Baker_2023	1
	32	1732.3	2.03	1.0	$NH_3/N_2/O_2/H_2/Ar$	0.5	77.4	J.Baker_2023	1
	33	1608.2	2.03	1.0	$NH_3/N_2/O_2/H_2/Ar$	0.5	77.4	J.Baker_2023	1
	34	1774.1	2.03	1.0	$NH_3/N_2/O_2/H_2/Ar$	0.5	77.4	J.Baker_2023	1

Source: Author (2024).

Table 12 – Experimental conditions for cluster 3.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
●	35	1252.2	1.27	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	7
	36	1022.7	1.21	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	7
	37	1409.1	1.1	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	7
	38	1144.6	1.28	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	7
	39	1600.6	1.19	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	7

Source: Author (2024).

Table 13 – Experimental conditions for cluster 4.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
●	75	1137.0	9.95	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	8
	70	1052.2	9.99	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	8
	71	1088.3	10.09	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	8
	72	1177.4	10.05	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	8
	73	1124.5	9.83	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	8
	74	1191.9	9.93	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	8
	76	1296.9	9.35	1.0	$NH_3/O_2/H_2/Ar$	0.7	29.77	J.Chen_2021	8

Source: Author (2024).

Table 14 – Experimental conditions for cluster 5.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
●	82	1437.2	1.21	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	5
	77	1199.4	1.19	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	5
	78	1265.5	1.2	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	5
	79	1516.7	1.2	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	5
	80	1335.6	1.23	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	5
	81	1658.9	1.19	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	5
	83	1825.8	1.22	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	5
	86	1704.0	10.17	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6
	88	1868.9	12.69	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6
	89	1551.0	9.44	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6
	90	1939.8	12.77	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6

Source: Author (2024).

Table 15 – Experimental conditions for cluster 6.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
●	84	1265.4	9.87	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6
	85	1316.9	10.12	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6
	87	1451.9	10.81	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6
	91	1206.6	9.99	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6
	92	1234.0	10.18	1.0	$NH_3/O_2/H_2/Ar$	0.3	28.74	J.Chen_2021	6

Source: Author (2024).

Table 16 – Experimental conditions for cluster 7.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
■	115	1488.0	14.49	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	43	1403.4	10.2	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	48	1448.4	10.2	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	49	1394.9	10.11	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	50	1444.2	9.82	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	51	1465.7	9.45	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	52	1441.0	10.52	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	2
	53	1505.0	10.42	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	2
	55	1522.3	10.55	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	2
	112	1456.0	14.69	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	113	1471.0	15.5	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	119	1433.0	15.3	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	124	1524.0	10.34	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	126	1617.0	10.03	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	135	1360.0	10.84	1.0	$NH_3/O_2/Ar$	0.0	16.71	Peng_2023	4
	137	1421.0	11.04	1.0	$NH_3/O_2/Ar$	0.0	16.71	Peng_2023	4
	139	1495.0	11.15	1.0	$NH_3/O_2/Ar$	0.0	16.71	Peng_2023	4
	155	1413.0	11.04	1.0	$NH_3/O_2/Ar$	0.0	24.33	Peng_2023	3
	156	1542.0	10.54	1.0	$NH_3/O_2/Ar$	0.0	24.33	Peng_2023	3
	157	1433.0	10.74	1.0	$NH_3/O_2/Ar$	0.0	24.33	Peng_2023	3
	159	1485.0	10.94	1.0	$NH_3/O_2/Ar$	0.0	24.33	Peng_2023	3
	183	1625.0	29.08	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	185	1593.0	29.59	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	186	1652.0	28.98	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	187	1564.0	30.09	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	226	1665.0	30.09	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	7
	228	1579.0	29.49	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	7
	230	1566.0	30.09	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	7
	242	1652.0	28.98	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4
	244	1619.0	29.18	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4
	245	1672.0	29.99	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4
	246	1586.0	29.49	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4

Source: Author (2024).

Table 17 – Experimental conditions for cluster 8.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
■	118	1731.0	12.87	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	40	1522.9	9.87	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	41	1608.9	9.37	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	42	1741.0	11.51	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	44	1653.2	9.96	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	45	1810.2	10.49	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	46	1675.21	10.22	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	47	1956.9	10.71	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	4
	54	1694.7	10.13	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	2
	56	1762.0	10.67	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	2
	57	1594.4	10.08	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	2
	58	1787.3	10.02	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	2
	59	1695.9	1.21	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	60	1733.6	1.14	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	61	1783.3	1.19	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	62	1613.4	1.02	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	63	1757.3	1.2	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	64	1804.8	1.14	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	65	1777.5	1.11	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	67	1634.3	1.27	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	68	1613.7	1.13	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	69	1667.7	1.24	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	93	1531.9	1.18	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	94	1625.8	1.22	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	96	1729.7	1.2	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	98	1757.8	1.24	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	100	1499.8	1.12	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	106	1645.0	12.06	0.99	$NH_3/O_2/Ar$	0.0	113.95	Peng_2023	1
	108	1646.0	11.04	0.99	$NH_3/O_2/Ar$	0.0	113.95	Peng_2023	1
	110	1760.0	11.04	0.99	$NH_3/O_2/Ar$	0.0	113.95	Peng_2023	1
	114	1621.0	12.67	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	116	1674.0	15.81	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	117	1601.0	13.07	1.0	$NH_3/O_2/Ar$	0.0	42.11	Peng_2023	2
	136	1659.0	12.36	1.0	$NH_3/O_2/Ar$	0.0	16.71	Peng_2023	4
	138	1750.0	12.26	1.0	$NH_3/O_2/Ar$	0.0	16.71	Peng_2023	4
	140	1930.0	13.78	1.0	$NH_3/O_2/Ar$	0.0	16.71	Peng_2023	4
	158	1712.0	10.54	1.0	$NH_3/O_2/Ar$	0.0	24.33	Peng_2023	3
	160	1872.0	13.98	1.0	$NH_3/O_2/Ar$	0.0	24.33	Peng_2023	3
	203	1686.0	11.04	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	205	1663.0	11.35	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	219	1768.0	10.64	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	9
	221	1719.0	11.04	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	9
	222	1776.0	10.23	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	9
	223	1618.0	11.04	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	9
	224	1866.0	28.37	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	7
	225	1810.0	28.57	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	7
	227	1723.0	28.37	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	7
	229	1716.0	29.59	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	7
	240	1808.0	28.37	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4
	241	1750.0	29.18	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4
	243	1742.0	29.59	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4

Source: Author (2024).

Table 18 – Experimental conditions for cluster 9.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
■	144	1331.0	21.08	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	8
	5	1437.0	40.0	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	4
	6	1581.0	42.0	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	4
	7	1342.0	41.0	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	4
	8	1484.0	38.0	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	4
	9	1280.0	42.2	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	4
	20	1303.0	20.5	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	3
	21	1525.0	19.5	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	3
	22	1294.0	20.0	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	3
	23	1405.0	20.0	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	3
	24	1264.0	20.6	1.0	$NH_3/N_2/O_2$	0.0	3.76	B.Shu	3
	120	1299.0	11.55	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	121	1368.0	11.15	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	122	1470.0	12.16	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	123	1373.0	11.75	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	125	1464.0	12.16	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	127	1266.0	10.74	1.0	$NH_3/O_2/Ar$	0.0	11.0	Peng_2023	5
	128	1228.0	13.17	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	6
	129	1266.0	12.46	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	6
	130	1304.0	11.65	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	6
	131	1281.0	13.48	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	6
	132	1344.0	11.55	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	6
	133	1300.0	12.26	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	6
	134	1367.0	12.26	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	6
	141	1380.0	21.08	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	8
	142	1275.0	21.08	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	8
	143	1422.0	21.08	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	8
	145	1462.0	19.56	1.0	$NH_3/O_2/Ar$	0.0	6.56	Peng_2023	8
	146	1180.0	12.97	1.0	$NH_3/O_2/Ar$	0.0	3.76	Peng_2023	10
	147	1295.0	12.26	1.0	$NH_3/O_2/Ar$	0.0	3.76	Peng_2023	10
	148	1200.0	12.77	1.0	$NH_3/O_2/Ar$	0.0	3.76	Peng_2023	10
	149	1364.0	12.36	1.0	$NH_3/O_2/Ar$	0.0	3.76	Peng_2023	10
	150	1266.0	11.55	1.0	$NH_3/O_2/Ar$	0.0	3.76	Peng_2023	10
	151	1414.0	11.45	1.0	$NH_3/O_2/Ar$	0.0	3.76	Peng_2023	10

Source: Author (2024).

Table 19 – Experimental conditions for cluster 10.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
■	166	2161.0	1.46	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	161	2330.0	1.39	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	162	2265.0	1.37	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	163	2073.0	1.51	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	164	2230.0	1.39	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	165	2000.0	1.48	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	167	1992.0	1.51	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	168	2404.0	1.39	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	169	2361.0	1.39	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	11
	263	2038.0	10.23	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12

Source: Author (2024).

Table 20 – Experimental conditions for cluster 11.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
■	175	1969.0	1.59	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	172	2008.0	1.55	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	174	2077.0	1.5	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	176	2023.0	1.48	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	177	1927.0	1.6	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	231	1923.0	11.04	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3
	238	2015.0	10.54	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3
	239	1984.0	11.04	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3

Source: Author (2024).

Table 21 – Experimental conditions for cluster 12.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
■	178	2312.0	1.45	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	170	2175.0	1.48	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	171	2164.0	1.46	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	173	2479.0	1.31	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	179	2389.0	1.36	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2
	180	2266.0	1.48	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	2

Source: Author (2024).

Table 22 – Experimental conditions for cluster 13.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
■	190	2253.0	1.43	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	189	2304.0	1.37	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	191	2087.0	1.4	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	192	2168.0	1.46	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	194	2091.0	1.44	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	196	2418.0	1.32	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	197	2489.0	1.31	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	198	2372.0	1.36	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	208	2253.0	1.35	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	209	2179.0	1.4	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	215	2454.0	1.3	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	216	2290.0	1.27	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8

Source: Author (2024).

Table 23 – Experimental conditions for cluster 14.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
▲	212	1895.0	1.45	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	66	1899.3	1.33	1.0	$NH_3/O_2/Ar$	0.0	28.15	J.Chen_2021	1
	95	1789.2	1.25	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	97	1820.8	1.21	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	99	1939.0	1.25	1.0	$NH_3/O_2/H_2/Ar$	0.05	28.23	J.Chen_2021	3
	107	1823.0	11.15	0.99	$NH_3/O_2/Ar$	0.0	113.95	Peng_2023	1
	109	1853.0	11.35	0.99	$NH_3/O_2/Ar$	0.0	113.95	Peng_2023	1
	111	1941.0	10.84	0.99	$NH_3/O_2/Ar$	0.0	113.95	Peng_2023	1
	193	2030.0	1.47	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	195	1987.0	1.53	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	5
	199	1947.0	10.74	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	200	1910.0	10.74	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	201	1742.0	11.15	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	202	1824.0	10.94	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	204	1790.0	10.94	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	206	2079.0	10.74	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	207	2005.0	10.94	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	6
	210	1927.0	1.49	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	211	2076.0	1.43	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	213	2029.0	1.44	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	214	1827.0	1.54	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	8
	217	2085.0	10.64	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	9
	218	2000.0	10.94	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	9
	220	1890.0	10.94	1.0	$NH_3/O_2/Ar$	0.0	114.35	O.Mathieu	9
	247	1926.0	28.27	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4
	248	1860.0	28.07	1.0	$NH_3/O_2/Ar$	0.0	231.04	O.Mathieu	4

Source: Author (2024).

Table 24 – Experimental conditions for cluster 15.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
▲	236	1742.0	11.15	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3
	181	1832.0	28.37	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	182	1771.0	28.78	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	184	1717.0	29.08	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	188	1894.0	28.07	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	1
	232	1838.0	10.74	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3
	233	1706.0	11.15	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3
	234	1801.0	11.04	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3
	235	1665.0	11.25	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3
	237	1626.0	11.55	0.5	$NH_3/O_2/Ar$	0.0	165.0	O.Mathieu	3

Source: Author (2024).

Table 25 – Experimental conditions for cluster 16.

Cluster	Index	Temperature [K]	Pressure [bar]	Φ	Mixture	H_2 Content	Dilution	Author	Data Set
▲	250	1844.0	28.88	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	10
	249	1906.0	27.86	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	10
	251	1660.0	29.49	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	10
	252	1762.0	29.59	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	10
	253	1599.0	29.28	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	10
	254	1676.0	30.8	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	10
	255	1584.0	28.98	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	10
	256	1962.0	10.64	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12
	257	1961.0	11.04	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12
	258	1750.0	10.84	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12
	259	1890.0	10.64	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12
	260	1729.0	11.15	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12
	261	1812.0	10.84	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12
	262	1651.0	11.45	2.0	$NH_3/O_2/Ar$	0.0	363.04	O.Mathieu	12

Source: Author (2024).