A Generalized Novel Group Contribution Kinetic Modeling Approach to Linear Alkene Metathesis

Nikhiel Isaiah Bansi

216013284

Submitted in fulfillment of the requirements for the degree of Master of Science in Chemical Engineering

College of Agriculture, Engineering and Science

University of Kwa-Zulu Natal

Supervisor: Professor David Lokhat

October 2023

As the candidate's supervisor, I Professor David Lokhat, have approved this thesis for submission,

Professor David Lokhat

Declaration

I, Nikhiel Isaiah Bansi declare that,

(i) The research reported in this dissertation/thesis, except where otherwise indicated, is my original work.

(ii) This dissertation/thesis has not been submitted for any degree or examination at any other university.

(iii) This dissertation/thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

(iv) This dissertation/thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.

(v) Where I have reproduced a publication of which I am an author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

(vi) This dissertation/thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation/thesis and in the References sections.

_____

Nikhiel Isaiah Bansi

Abstract

Traditional modeling approaches for linear metathesis systems involve the use of complicated approaches – This paper proposes a generalized novel kinetic modeling technique, in which the substituent (alkyl) groups on the transition states are postulated to have a unique effect on the rate at which the controlling step (transition state dissociation) occurs.

Metathesis involves the redistribution of the *carbon-carbon double bond* across constituents, to create longer chain hydrocarbons, which significantly enhances product value in an industrial context. This work proceeds by unpacking the mechanics of the 1-Hexene (linear) metathesis pathway and utilizes this as a basis for developing a group contribution kinetic modeling approach. The substituent (alkyl) groups present in the forming transition state complexes were used to define (identify) rate constants that would potentially control the rate of formation of the various transition state complexes. A key assumption in this work was that the largest externally attaching substituent, and the substituent existing in the metallacyclobutane complex, be selected as the groups that define the formation of the respective transition state complexes. These observations resulted in a system of 25 uniquely defined (identified) rate parameters that were derived based on the *defining groups* observed within the mechanisms and are sufficient to describe the metathesis system under investigation.

Literature experimental data, at a range of temperatures (420℃ – 460℃) was readily available and was used for the purpose of fitting the identified rate parameters, by simultaneously solving the system of differential equations that result from this system. Given the system size (25 parameters), and complexity of the interactions in the system, evolutionary and swarm optimization techniques were found to be fit for this purpose. It was found that combination between a *genetic algorithm* (GA) and *particle swarm optimization* (PSO) approach yielded identified parameters that minimized the overall error of the prediction.

Kinetic parameters were identified, and Arrhenius plots were developed – These allowed for the activation energies (E_a) and pre-exponential factors (A_o) to be determined for each parameter. This was tested using literature experimental datasets, and predictions were found to present with a *modified fitness* value between 1.968 and 50.55. This illustrates that the novel group contribution kinetic modeling approach was suitable to define the interactions in the system. Further research is required to generalize this result over other alkene metathesis systems; however, this work proves that the approach is viable.

Extending this work to *ring closing metathesis* (RCM), is a future area of research that is of particular importance, as pharmaceutical intermediates result from these processes, and represents the unique opportunity to create a synergistic industrial landscape.

Contents

List of Figures

List of Tables

1. Introduction

Linear alkene metathesis is a chemical reaction, involving carbon–carbon double bond ($R_1C{=}CR_2$ & $R_3C{=}CR_4$) dissociation ($R_1C{-}{-}CR_2$ & $R_3C{-}{-}CR_4$), constituent redistribution (R_1, R_2, R_3 & R_4) and bond formation ($R_1C{=}CR_3$ & $R_3C{=}CR_2$) in alkene molecules (Lokhat *et al.*, 2008). Redistribution involves switching the constituents present on either side of the double bonds, across reactant species; such that unique species form.

Typical alkene effluent streams from synthetic fuel processing plants, consist of carbon chains with a length of C_4 to C_8 (Prithipal *et al.*, 2015). According to Prithipal *et al.* (2015), these carbon chain lengths are not suitable to hydroformylation or alkylation; therefore, longer carbon chains (C_{10}-C_{16}) are preferred. Hydroformylation and alkylation would result in the production of detergent range alkenes, which would create new opportunities for this industry in South Africa. Furthermore, low carbon number alkenes are usually more volatile than high carbon number alkenes (Clark and Nguyen, 2020); therefore, the ability to transform variable composition streams would reduce (C_4-C_8) alkene storage costs, enhance process safety (reduced volatile presence), and produce a valuable product (C_{10}-C_{16} detergent range alkenes).

Kinetic modeling forms the basis for the design and scale-up of the alkene metathesis process, and kinetic studies are not widely available in literature; as such, the design of these processes rely on existing unit performance data or lumped kinetic models (Froment, 2005). Kinetic models describing this process are of interest to SASOL, and the generalised petroleum industry. The variable feedstock to alkene metathesis processes generally consists of *terminal linear alkenes*, available from the SASOL coal to liquid process and various other petroleum process effluents (Prithipal *et al.*, 2015). The terminal alkenes occur in low carbon chain lengths (C_4 to C_8); as such, alkene metathesis is leveraged to increase the length of these carbon chains. This enhances the value and usefulness of these alkenes by creating a synergistic landscape with industries such as the pharmaceutical and other chemical sectors (Prithipal *et al.*, 2015).

The issue that arises is due to the composition of the alkene metathesis feed stream. Metathesis processes are unique and would require the development of kinetic models for those specific feedstock compositions and system configurations. *Lumped kinetic* (LK) approaches involve parameters that depend on the feed composition and reactor configuration, which renders them inflexible for different operational set-ups (Svoboda *et al.*, 1995). Therefore, the development of a generalised kinetic model (or modeling technique) that can adequately predict reactive interactions, without the need to develop system specific kinetics, would be more robust.

This work aims to address this need by creating a modeling methodology specific to the type of substituents (alkyl groups) that are involved in the metathesis (redistribution) process. A methodology that is specific to the type of groups could be generalised across different system configurations (feedstocks & reactor set-ups), that employ the same catalyst (8wt% WO_3/SiO_2), provided that the chemistry is understood.

Alkene metathesis product compositions are highly variable, due to isomerization and secondary reactions that can occur on the acidic catalyst surface (Lokhat *et al.*, 2008). This is an indication of a complex reaction network, which cannot easily be replicated using a traditional LK approach (Svoboda *et al.*, 1995). Lumped models become more complex as the number of components/ interactions in the system increases; therefore, a novel group contribution modeling methodology is proposed.

This work suggests that there is a unique effect associated with the rate of attachment of different alkyl groups, during the formation of the reaction specific transition states. This would result in a reaction specific effect that can be defined using a certain combination of the alkyl groups in the kinetic model. This method of modeling should allow for sufficient comparability to the chemistry of the metathesis reaction, which will allow for the treatment of complex reaction networks and interactions.

The *single event* (SE) kinetic modeling approach is an attractive alternative to the proposed technique, as it would provide a more simplified representation of the kinetic network in complex systems, without the computational intensity of the LK approach (Svoboda *et al.*, 1995). The SE approach essentially involves fitting (intermediate) rate coefficients, to an elementary step in a reaction, using product distribution data. This translates to fewer rate coefficients, as the number of elementary steps will always be less than the number of molecules (Froment, 2005). SE rate coefficients are modelled through statistical thermodynamics and transition state theory, which will ensure independency of the feed composition. The SE method incorporates *tuning parameters* which account for the specifics of the type of catalyst used. The limitation of the SE approach is that it oversimplifies complex systems, by assuming that each elementary step (event) occurs in isolation, and it becomes difficult to account for the effect of side reactions (isomerisation) that usually occur in metathesis systems (Froment, 2005).

Froment (2005) also highlights other simulation/ modeling methods, namely *Monte-Carlo simulation*, and *Structure Oriented Lumping*; however, these are based on the molecular structure of the feedstock. Therefore, there would be a significantly larger number of rate parameters associated with these modeling approaches.

The proposed group contribution technique is also based upon *transition state theory*, but this work involves a simplification in accounting for the dissociation of the formed transition states. This theoretically-based simplification estimates transition state dissociation, without the need for complex statistical thermodynamics.

Optimisation techniques are gaining increasing application in the field of engineering, and artificial intelligence. Therefore, similar techniques will be used in this work to fit the identified rate parameters and assess the performance of the proposed group contribution method.

The versatility of the metathesis reaction also allows for *greener chemical synthesis*, when compared to traditional methods, in industries such as the pharmaceutical and fine chemical industry. Traditional synthesis pathways usually consist of unavoidable – sometimes inefficient – pathways to form the desired products. However, the 'shifting nature' of the metathesis mechanism makes it possible to form the required intermediates

using milder (less energy intensive) conditions, in the presence of a suitable catalyst. Pharmaceutical intermediates can be formed by the *ring closing metathesis* (RCM) of linear alpha-olefins, via small scale processes (Dos Santos *et al.,* 2018). Therefore, this work provides a framework for the creation of a circular economy, by creating a connection between the petrochemical and pharmaceutical sectors. An understanding of the mechanism of RCM, and its relation to linear metathesis, would allow for the design and optimization of reactors aimed at minimizing undesired reaction occurrence, and the implementation of suitable control systems to ensure product quality (Fandrick *et al.,* 2014). This would allow for the development of novel pharmaceuticals, improved synthesis methods, and possible operation on a larger industrial scale.

Aside from petroleum refineries, bio-refineries can also be used to produce alpha-olefins (alkenes) from plant-based oils (Dos Santos *et al.*, 2018). According to Cchange (2020), valuable products, such as polymers, lubricants, and surfactants, can be produced from these (petroleum & bio refinery) alpha-olefins; therefore, an improved understanding of the linear metathesis mechanism would directly promote the bio-refinery (greener) concept, with the possibility of product upgrading.

Aim

- To develop and assess the performance of a generalized novel group contribution kinetic modeling approach for linear alkene metathesis.

Objectives

- To apply various techniques to develop a linear alkene metathesis (1-Hexene metathesis) kinetic model, based on various literature supported assumptions, novel modeling techniques, and proposed ideas, to effectively convert variable composition (alkene) streams, into higher value products, and promote the circular economy.
- To apply intelligent optimization techniques, using the available literature (1-Hexene) metathesis data, as a means of developing and validating a potential group contribution kinetic model.

2. Literature Review

2.1. Metathesis Mechanism Overview

The olefin metathesis process is catalyzed by metal oxide catalysts, which undergo transformation to form *metal alkylidenes* during the reaction (Schaller, 2020). *Alkylidenes* are active for the metathesis process and form a cyclic complex, known as a *metallacyclobutane*, during reaction with the olefins. A *metallacyclobutane* is a four-member ring which forms and readily dissociates during the metathesis process, to create a *new metal alkylidene* and an alkene product (Schaller, 2020). The new *metal alkylidene* then reacts with another alkene, to form the other alkene product and another *alkylidene*. Therefore, the overall metathesis process consists of various steps, in which the *alkylidene* created in one reaction will produce the desired product (long chain internal alkenes) in the next reaction step. An important fact is that the *new alkylidene* always forms with the remaining constituents of the *metallacyclobutane* ring, and it never contains what was a constituent of the *original alkylidene*.

2.1.1. The Chauvin Mechanism

The Chauvin mechanism describes the metathesis process, including the formation and dissociation of the *metallacyclobutane* complex (Schaller, 2020). The metallacyclobutane complexes are effectively transition state complexes that can dissociate to yield either the reactants or products. A typical representation of this overall process in provided in Figures 2-1 and 2-2.

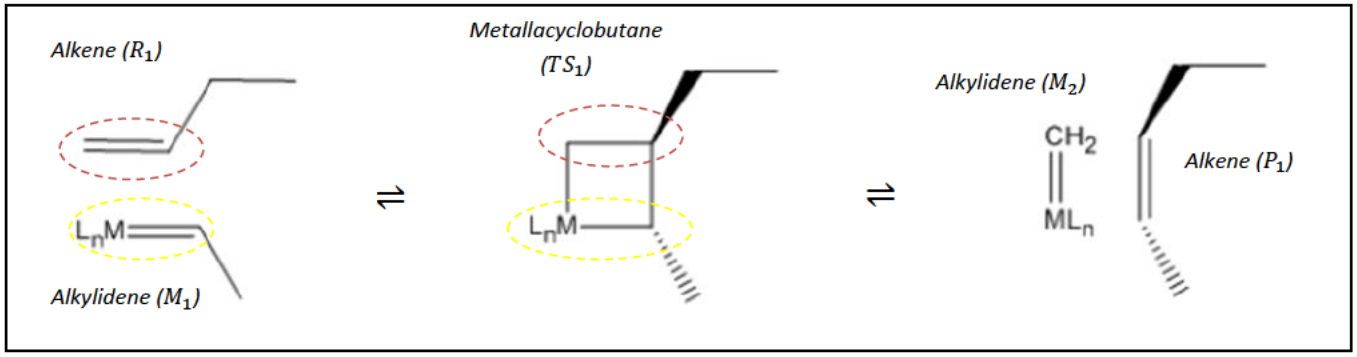


Figure 2-1. Chauvin Mechanism Process [1/2] (Schaller, 2020)

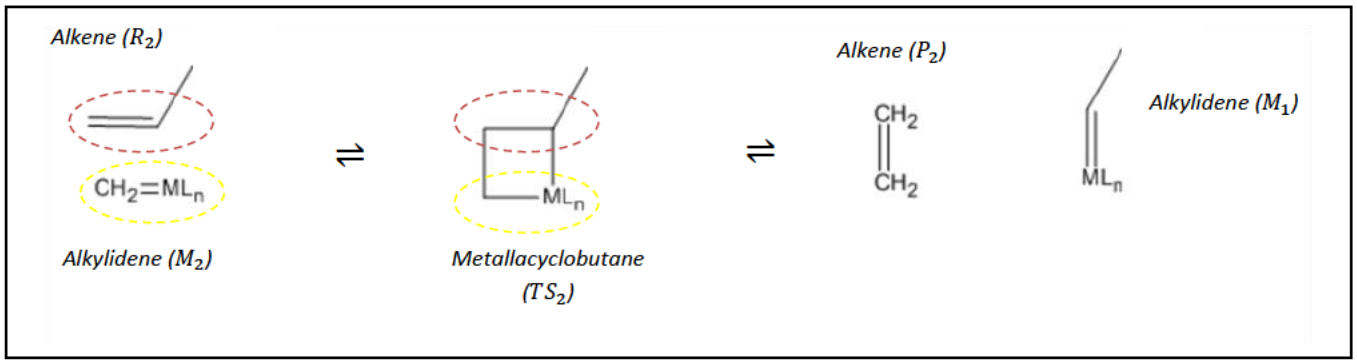


Figure 2-2. Chauvin Mechanism Process [2/2] (Schaller, 2020)

The mechanism in Figure 2-1 describes the first elementary step in a metathesis cycle – In this step, reactant R_1 reacts with alkylidene M_1 to form transition state TS_1. TS_1 can then dissociate to form the product P_1, or reform the reactant R_1. The direction of dissociation is determined by the stability of the bonds made to form a transition state and is directly driven by the chemical equilibrium of the process. To complete a metathesis cycle, the formed alkylidene M_2 participates in the elementary process of Figure 2-2. The process follows along resulting in the reaction of R_2 with alkylidene M_2, formation of transition state TS_2, the dissociation of transition state TS_2 and the formation of product P_2 and alkylidene M_1. This indicates that there is no defined order of the steps, instead, the order is driven by the availability of the required alkylidene and reactant. From the mechanisms in Figure 2-1 and 2-2, longer chain alkylidenes (>1 carbons) will form *internal* alkene products, from *terminal* alkene reactants (Schaller, 2020).

2.1.2. 1-Hexene Self-Metathesis

Experimental data available in literature was used for the purpose of this work. Prithipal *et al.* (2015) and *Lokhat et al. (2008)* investigated the metathesis of 1-hexene over an 8wt% WO_3/SiO_2 catalyst, in a gas-phase fixed bed reactor at 420°C to 500°C. The summarized results of these experiments are provided in Table A-1, and Table A-2, of Appendix A, and it will be used to investigate the existence of a group contribution approach to the kinetic modeling of alkene metathesis. Prithipal *et al.* (2015) also investigated the effect of doping the 8wt% WO_3/SiO_2 catalyst with alkali metal (potassium) ions and found that this decreased the Bronsted acidity of the catalyst sites, which promoted metathesis and isomerization action over the cracking action. However, over doping with the alkali metal ions was also found to block adjacent Lewis acid sites, reducing the metathesis activity due to steric factors; therefore, an optimal alkali metal doping exists (Prithipal *et al.*, 2015).

For the metathesis process to occur, a *carbenoid chain initiator* (metal carbene/ alkylidene) must form. The formation of the carbenoid chain initiator for the self-metathesis of 1-hexene was illustrated by Prithipal *et al.* (2015), and the resulting initiator complex is shown in Figure 2-3. The initiator complex forms during the induction period, during which the catalyst activity gradually increases until a stable value is reached (Prithipal *et al.*, 2015). In Figure 2-3 and in this work, the *M* is used to denote the tungsten complex in the red dashed border. The carbenoid chain initiator participates in the initiator step shown in Figure 2-3, to produce one of the alkylidenes required for the metathesis process. Both the alkylidenes (M_1 or M_2) can form and participate in the 1-Hexene self-metathesis mechanism of Figure 2-4. This implies that there is no specific order for the elementary steps shown in Figure 2-4. A further note about the initiator is that it only exists for a limited number of metathesis cycles, after which the chain of reactions will terminate (Prithipal *et al.*, 2015).

The metathesis process of Figure 2-4 begins with the cleavage of the alkene double bond in 1-hexene and the alkylidene (M_1), followed by the simultaneous coupling of these species. This process forms weak bonds to create the first metallacyclobutane complex (TS_1), which is a puckered ring that contains a dihedral ring angle in the range 10° to 30° (Prithipal *et al.*, 2015). This complex then decomposes to form the first product (5-

decene) and the alkylidene required (M_2) in the next elementary step. The same alkene double bond cleavage and bond formation then occurs with 1-hexene and the second alkylidene (M_2), which forms a second metallacyclobutane intermediate (TS_2). The second cyclic intermediate (TS_2) then decomposes to form the original alkylidene (M_1) and the second product (1-ethene or ethylene).

The molecularity of the 1-hexene self-metathesis process of Figure 2-4, indicates that two molecules of 1-hexene produce a single molecule of ethylene and 5-decene, in a single cycle. All the previously mentioned steps occur at the surface of the catalyst; hence it considered a surface reaction.

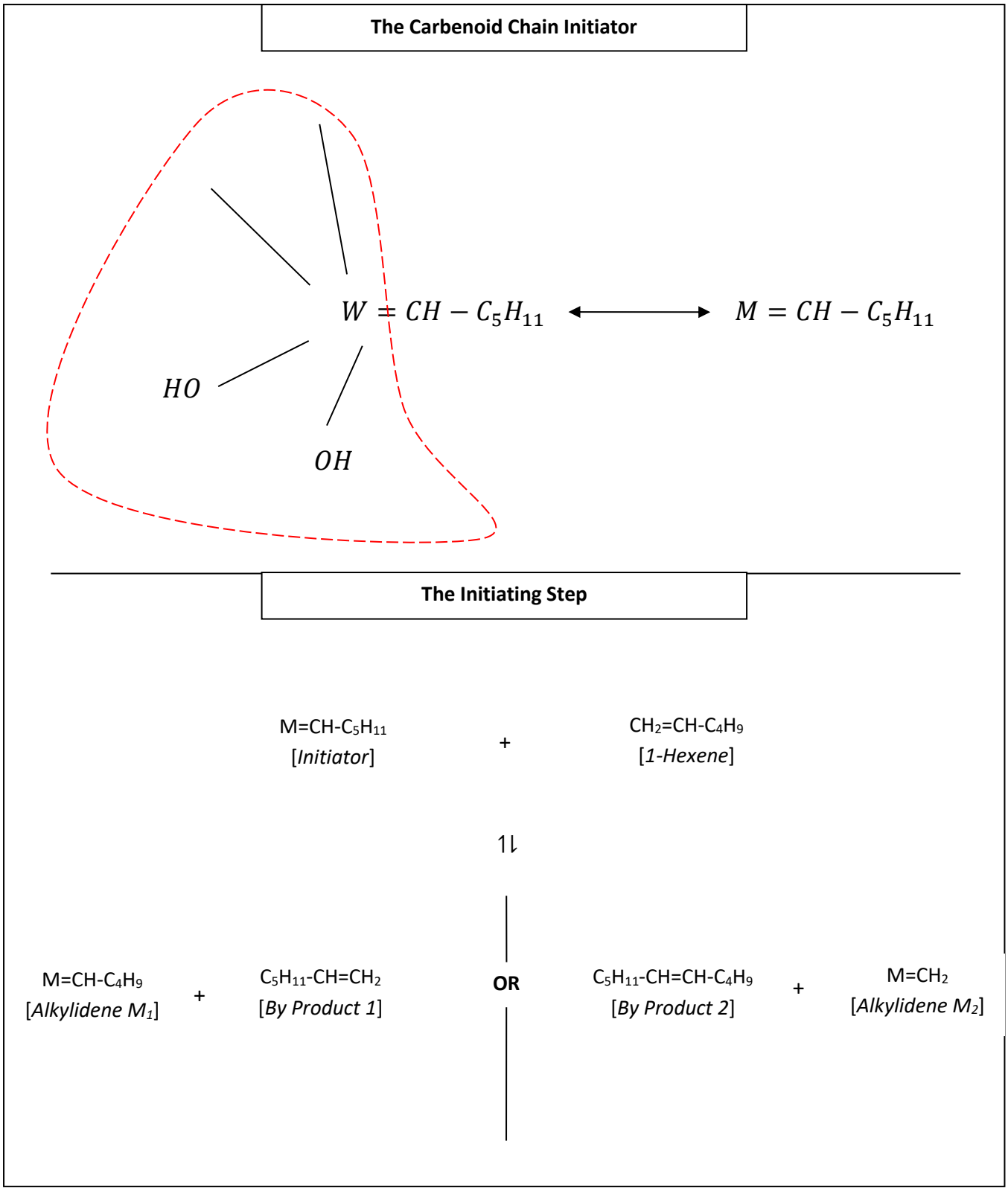


Figure 2-3. Carbenoid Chain Initiator & Initiating Step Illustration (Prithipal *et al.*, 2015)

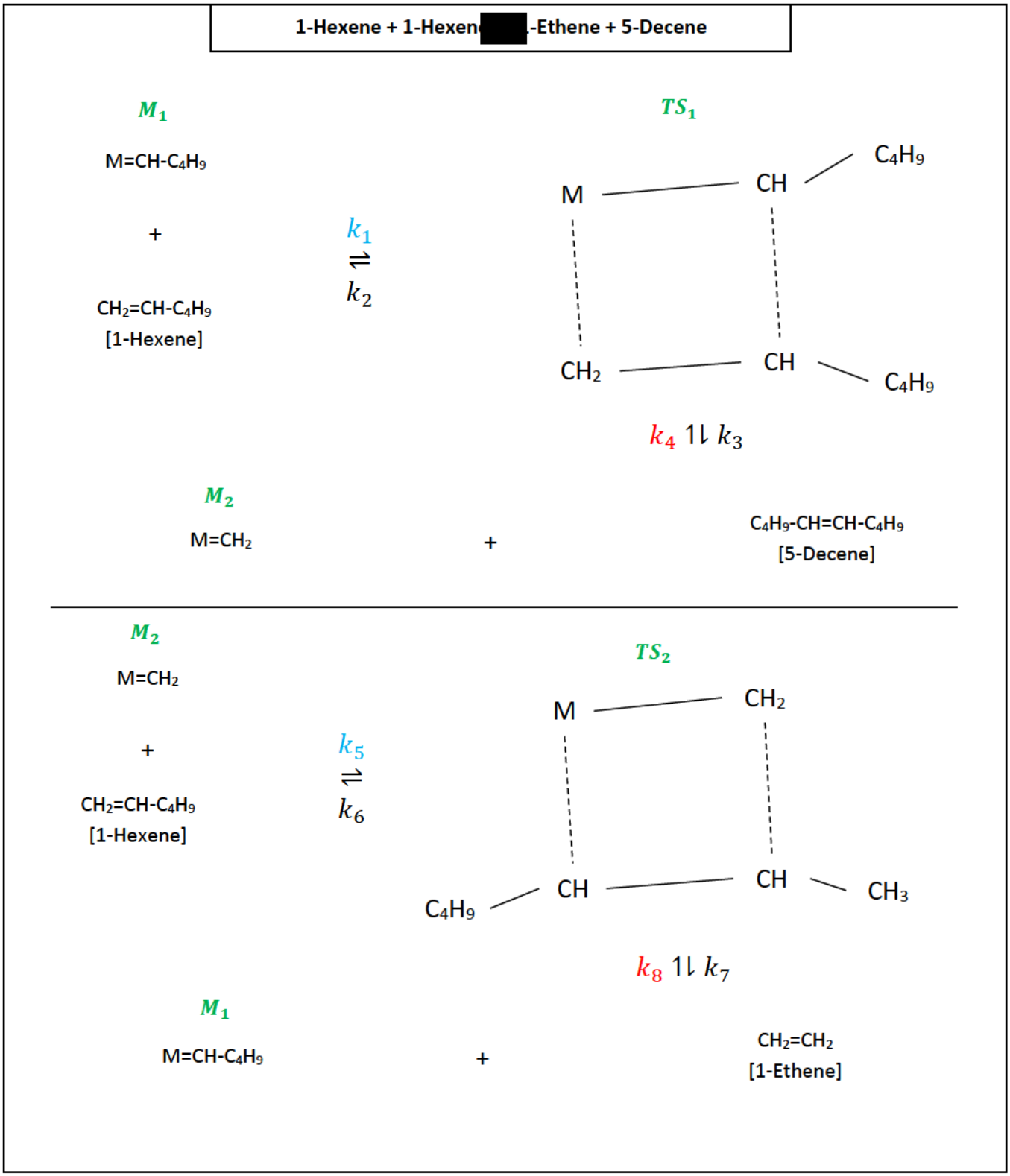


Figure 2-4. 1-Hexene Self-Metathesis Cycle (Prithipal *et al.*, 2015)

2.1.3. Stability of The Metallacyclobutane (Transition State) Complexes

Considering that the *metathesis* reaction occurs by transition state dissociation and formation, which is a well understood theoretical mechanism, it becomes important to understand how the alkyl groups influence the stereoselectivity of the products that result from the metallacyclobutane complex, and the resulting kinetics of the metathesis reaction. Kapteijn and Mol (1981), investigated the selective formation of geometrical product isomers in the metathesis of both propene and pent-2-ene (cis and trans), over several solid catalysts (WO_3/SiO_2, etc).

There are three distinct kinds of metathesis reactions, these are (Kapteijn and Mol, 1981),

- Productive metathesis - Yields new products.
- Degenerate (or non-productive) metathesis – Here the exchange of alkylidene groups does not result in new products.
- Cis-trans isomerisation – Involves an exchange in steric make-up, and the effect of this process is explained in Section 2.2.

According to Kapteijn and Mol (1981), the preference for the trans-isomer increases as the chain length and bulkiness of the alkyl groups in the product molecule increase; however, this trend is also very closely influenced by the type of catalyst used. Kapteijn and Mol (1981), observed that the type of alkyl groups present in the metallacyclobutane complex (intermediate) would directly influence the product isomers that form – This occurs as the intermediate does not exist in a flat ring structure, instead it exists in a puckered state, and it bent along the C_1-C_3 axis shown in Figure 2-5.

This indicates that repulsive interactions between the alkyl ring substituents, as well as the interactions of the substituents with the groups on the transition metal (or catalyst surface) govern the stereoselectivity of the reaction. Therefore, steric repulsions exist in the metallacyclobutane complex and will cause an increase in activation energy, with a resulting reduction in its formation rate (Kapteijn and Mol, 1981).

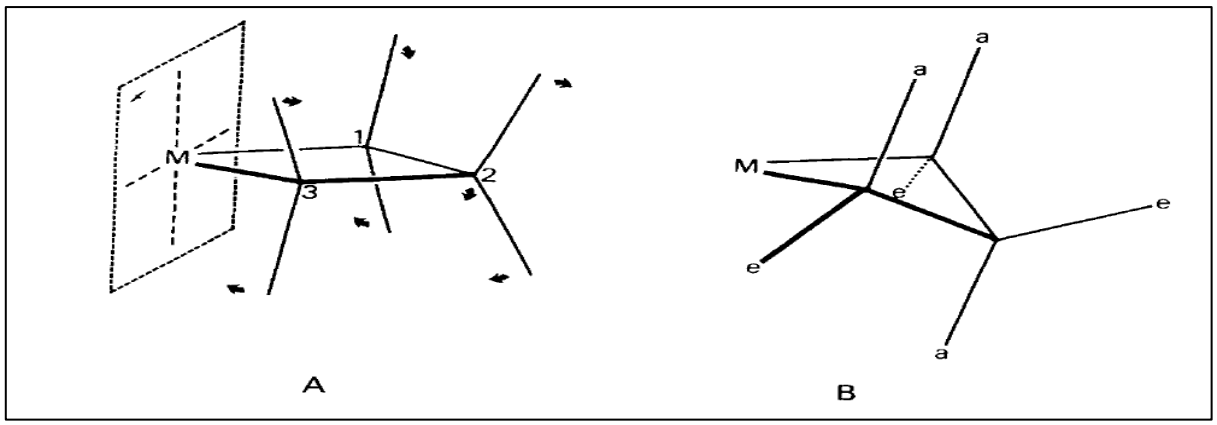


Figure 2-5. Structures of a Metallacycle Complex (A) Flat, (B) Puckered Ring

2.2. Metathesis Product Isomer Stability

There are two types of linear alkene metathesis processes that can occur, namely primary or secondary metathesis. Primary metathesis occurs when a terminal alkene reacts to produce an E/Z (trans/ cis) product, and Secondary metathesis occurs when the E/Z product further reacts to form internal alkenes. The type that occurs is dependent on the sterics/ spatial arrangement of the carbons in the reactants, and the selectivity toward a certain stereoisomer can be quantified in terms of the E/Z stereoselectivity ratio.

2.2.1. Cahn-Ingold-Prelog (CIP) priority rules

E/Z refers to the naming conventions for geometric isomers/ stereoisomers. According to the Cahn-Ingold-Prelog (CIP) priority rules (Hunt, 2020),

- An E-Alkene contains priority groups on opposite sides.
- A Z-Alkene contains priority groups on the same side.

Priority groups are assigned based on the atomic number of the (first point of difference) atom attached to a carbon involved in a double bond, where '1' indicates high and '2' indicates low priority (Hunt, 2020). The main criterion for assigning a priority is that the high priority (Hunt, 2020) is given to the constituent that has the highest atomic number. If there are two of the same atoms attached on one side of a double bond, then the atoms attached to these carbons must be listed in order of decreasing atomic number (Bruner, 2020), and each atom must be compared until a difference in atomic number is found.

2.2.2. Isomer Selectivity

The thermodynamic stability of alkene isomers controls the E/Z stereoselectivity of the metathesis reaction (Donde *et al.*, 2008). The E-isomer (trans) is usually more thermodynamically stable, whilst the Z-isomer (cis) is always more reactive (Donde *et al.*, 2008); therefore, at higher conversions the E/Z ratio will increase as more Z-isomers will be converted to E-isomers through secondary metathesis.

Secondary metathesis is noted to interconvert isomers to selectively form the more stable isomers (Donde *et al.*, 2008). However, the E/Z ratio is lower at reduced conversions as the primary metathesis process is much faster and is more selective to the more reactive Z-isomer (Donde *et al.*, 2008). This implies that complex chemical systems would not simply undergo self-metathesis, instead, the reaction would proceed via a number of metathesis pathways that can involve the products, until the initiator is no longer present.

2.3. Catalyst Types

There are four distinct generations of olefin (alkene) metathesis catalysts. These will be discussed in order of their discovery. An essential factor that determines the catalyst activity is the acidity of the active sites, i.e., either *Lewis acid* or *Bronsted acid*. A *Lewis acid* refers to a catalyst site that can accept an electron pair, whilst a *Bronsted acid* refers to a site that can donate a proton (H^+) (LibreTexts, 2019). *Bronsted acids* are generally stronger than *Lewis acids*. The relationships between Metathesis and *Green chemistry* will also be highlighted in Section 2.4, as it is related to the type of catalyst used.

2.3.1. Black Box Catalyst

These are heterogeneous catalysts that consist of a high valence transition metal halides, oxides, or oxo-halides, with an alkylating co-catalyst (Toreki, 2002). These catalysts are often used with silica/ alumina supports. According to Toreki (2002), these catalysts are highly active but have a low tolerance for functional groups, as they are acidic (*Lewis*) catalysts. Therefore, a very small percentage (1%) of the catalyst is considered active (Toreki, 2002).

Typical examples of this catalyst are $WCl_6/SnMe_4$ and $Re_2O_7Al_2O_3$, which are usually utilized in Ring opening metathesis processes (ROMP).

2.3.2. Tebbe's Reagent (Titanium based)

These are Titanium based catalysts that are well defined in terms of their function; therefore, it was used to confirm the *Chauvin mechanism* for olefin metathesis and perform metallacyclobutane studies (Grubbs, 2006). According to Grubbs (2006), it was concluded that metallacyclobutane complexes act as the olefin metathesis catalysts. The only mechanistic difference between this and later catalysts, was that the final step was not reversible (Toreki, 2002). Compared current generation metathesis catalysts, *Tebbe's reagent* is not tolerant of *carbonyl* functional groups, and undergoes *Wittig*-like reactions to form methylene derivatives from ketones, aldeydes and carbonyls (Toreki, 2002). Tebbe's reagent was found to be highly sensitive to air and is difficult to synthesize/ prepare (Grubbs, 2006).

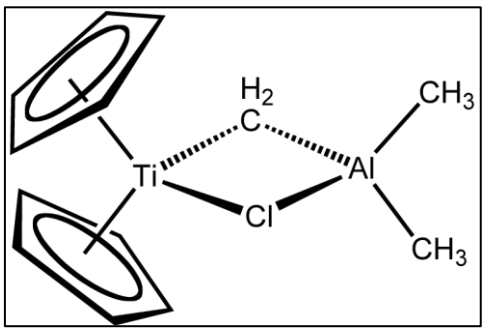


Figure 2-6. Tebbe's Reagent (Grubbs, 2006)

2.3.3. Schrock's Catalyst (Tungsten, Molybdenum or Rhenium based)

Molybdenum (Mo) and Tungsten (W) catalysts are highly efficient and active for olefin metathesis (Toreki, 2002). Molybdenum is also considered a *chiral catalyst*, which implies that *enantiomers* of products can form (Schrock and Hoveyda, 2003). This is particularly useful in the pharmaceutical industry, in which *chiral molecules/ enantiomers* can offer completely different pharmaceutical functions. The issues with these catalysts arise due to their *oxophilic* metal centers, i.e., they readily bond to oxygen atoms (Grubbs, 2006). Oxygen usually attaches to the metal center in the form of *alkoxide ligands*, which results in a poor functional group tolerance; thus, catalyst preparation under inert (O_2 & H_2O free) conditions is necessary. The higher operating temperature (420 to 500℃) observed with tungsten catalyst creates a resiliency to oxygen poisoning; however, these conditions are more favorable for *side reactions* (Isomerisation, Cracking and Secondary metathesis) (Prithipal *et al.*, 2015).

Since the active catalyst structure is well known, reactivity tuning can readily be achieved by adjusting the carbon chains (alkyl/ aryl groups in Figure 2-7) attached to the *alkoxide ligands* (Toreki, 2002). Bulkier *alkoxide ligands* will allow for more efficient olefin metathesis (Schrock and Hoveyda, 2003). The catalyst can be tuned to be active for other types of metathesis processes. Prithipal *et al.* (2015) performed a catalyst reactivity tuning study, in which they found that primary metathesis can be favored by optimally doping the catalyst with *alkali metal acid inhibitors* (potassium ions). This reduced the acidity of the *strong acid sites* present in the tungsten catalyst; thereby, reducing the selectivity toward the *side reactions* (Isomerisation, Cracking and Secondary metathesis). Over doping the catalyst (with potassium) was noted to reduce metathesis activity by blocking the Lewis acid sites responsible for *carbene/ alkylidene* formation. In general, the stronger Bronsted acid sites catalyze Isomerisation and Cracking reactions, whilst *Lewis acid sites* catalyze Metathesis and Isomerisation (Prithipal *et al.*, 2015).

Schrock's catalysts are very sensitive to air and water; however, ensuring that the reaction environment is free from these contaminants would result in a catalyst that is highly tolerant of *other* functional groups (Toreki, 2002). *Schrock's catalyst* is noted to be effective in the presence of *phosphanes* and *thioethers*, which would normally decompose Ruthenium complexes (Schrock and Hoveyda, 2003).

Tungsten-based (WO_3/SiO_2) metathesis offers many advantages over molybdenum and rhenium-based metathesis. The catalyst is tolerant to poisons (oxygenates) in the feed at high temperatures (400℃ to 600℃), it has a longer lifespan compared to molybdenum and rhenium, it has a relatively constant activity at high coke conditions and does not become brittle when regenerated (Lokhat *et al.*, 2008). The main disadvantage of combining tungsten-based catalysts with a high temperature process is that it catalyses the cracking and isomerisation of alkenes, leading to a more variable product yield (Lokhat *et al.*, 2008).

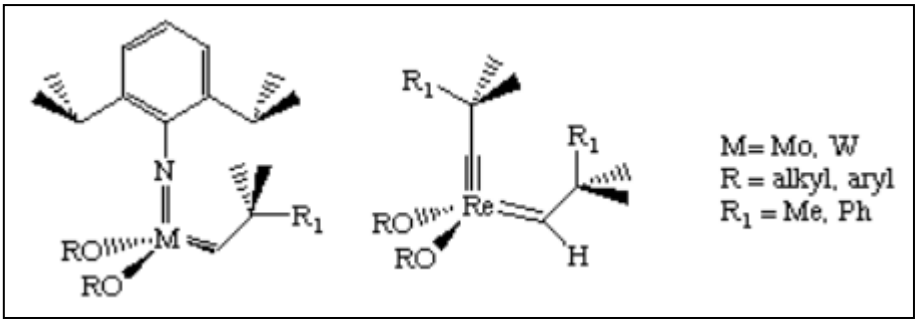


Figure 2-7. Typical Schrock's Catalyst Examples (Toreki, 2002)

2.3.4. Grubbs Catalyst (Ruthenium based)

Grubbs catalyst is highly tolerant to functional groups, which allows for metathesis even in the presence of water. However, this added functional group tolerance comes with the issue of low metathesis rates, when compared to *Schrock's catalyst* (Toreki, 2002). According to Grubbs (2006), only a small percentage of the (1st generation) ruthenium catalyst added formed an active catalytic center; furthermore, the structure of the active catalyst was undefined which resulted in difficulty with reactivity tuning/ catalyst optimization.

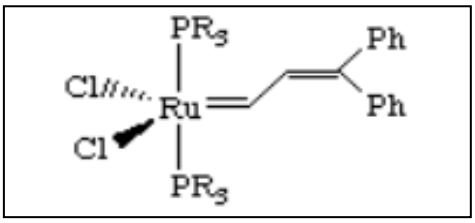


Figure 2-8. Unoptimized (1st Generation) Grubbs Catalyst (Toreki, 2002)

Further developments with this catalyst have indicated that a *Ruthenium (II) complex* and a strained olefin are necessary to form an active catalyst system (Grubbs, 2006). As a result of this discovery, the active catalyst structure was eventually determined, and is illustrated in Figure 2-8, which allowed for catalyst optimization to become possible. *Phosphine ligands* were noted to support and stabilize the transition metal complex, which is illustrated in Figure 2-8. Ruthenium is very rich in *d-electrons*; therefore, strongly *electron donating ligands* are required to ensure high catalyst activity.

Ruthenium catalysts have selectivity toward reactions involving *π-acid ligands* and *soft Lewis bases*, which make them fundamentally different from Molybdenum and Tungsten catalyst (Grubbs, 2006). *π-acid ligands*, which refer to alkenes for the purposes of this work, bond to the catalyst metal center according to the *Dewar-Chatt-Duncanson model*. This model is illustrated in Figure 2-9; in which, the *π-acid ligand* (alkene) donates electrons from its filled *π-orbitals* to an empty *d-orbital*, which occurs in synergism with filled *d-orbital* back-donation to empty π^**-antibonding orbitals*. Typical examples of ions that form *soft Lewis bases* are provided in Figure 2-10.

Hard Lewis acids	Borderline Lewis acids	Soft Lewis acids
H^+, Na^+, K^+ Be^{2+}, Mg^{2+}, Sr^{2+} Sc^{3+}, Fe^{3+}, Ln^{3+}	Fe^{2+}, Ni^{2+}, Zn^{2+}	Cu^+, Ag^+, Au^+ Cd^{2+}, Pd^{2+}, Pt^{2+} Tl^{3+}
Hard Lewis bases	**Borderline Lewis bases**	**Soft Lewis bases**
H_2O, NH_3, R_2O Cl^-, NO_3^-,	Br^-	RSH, R_2S I^-, CN^-

Figure 2-9. Dewar-Chatt-Duncanson model (Marshall, 2014)

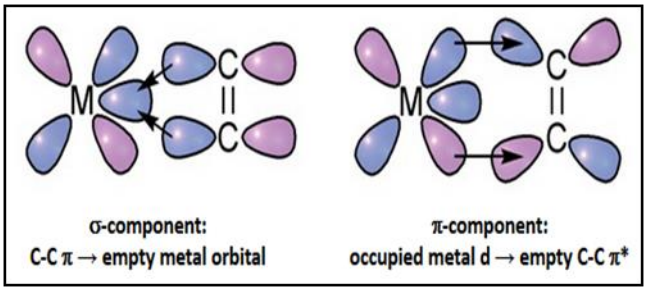


Figure 2-10. Soft/ Hard Lewis Acid/Bases (Lundberg, 2006)

2.4. Green Chemistry

Green chemistry refers to the overall methods of conducting reactions, with the basic focus on minimizing the impact on the environment, through more efficient chemical processes. There are three ways in which Metathesis could hypothetically satisfy the goals of green chemistry (Grubbs, 2006),

- It enables the discovery of *more efficient pathways* to a particular product, by avoiding/ reducing the formation of by-products.
- It allows for the use of *renewable resources*; reason being, Metathesis pathways are generally less energy intensive than traditional synthesis pathways. Therefore, current solar/ hydroelectric capacity would be sufficient to supply majority of the required energy. In this case, *Molybdenum* and *Rhenium based catalyst* would be preferred as they catalyze reactions at much lower temperature conditions, compared to a *Tungsten based catalyst* (Prithipal *et al.*, 2015).
- It promotes sustainable bioprocesses. Products formed by biochemical processes are generally variable and not at a useful carbon chain length. The introduction of well understood metathesis processes and kinetic models to these fields has the potential to improve the quality of fuels or biodegradable solvent alternatives.

In this work *Tungsten based catalyst* is used which functions in a higher temperature range, i.e. between 400°C to 600°C. Higher temperatures are more difficult to attain using renewable energy sources but can be achieved using more modern techniques that usually involve concentrated solar technologies, such as a parabolic dish or central receiver tower (Teske *et al.*, 2022). Typical industrial set-ups involve a solar concentrator to focus radiation, a solar receiver to absorb heat, a transfer or storage medium to facilitate heat movement through the process site, and a power conversion system or a transfer method to direct process heat to the intended system/ process (Teske *et al.*, 2022). Solar energy can readily be transferred to a medium, such as a synthetic oil, stored in the form of electricity for the creation of green hydrogen through renewable electrolysis, and used in other direct electrical heating methods.

2.5. Generalized Types of Reaction Mechanisms

Reaction mechanisms generally consist of multiple *elementary reactions* that occur as a combination of *series* or *parallel reactions*, depending on the species that are available or produced during the reaction (Levenspiel, 1999). The distinction between *series* and *parallel reactions* are illustrated by means of Figure 2-11. In accordance with Figure 2-11, *series reactions* require that the steps of a reaction occur in a specific sequence, to produce a certain species, e.g., *species S*. However, *parallel reactions* allow for the formation of products via paths that do not require a specific sequence of steps. *Parallel reactions* can take the form of *competitive reactions*, where a single species competitively reacts to form multiple products, or *side-by-side reactions*, where reactions occur independently of each other. These are important concepts to understand, as they have a direct impact on the development and understanding of a *reaction mechanism*.

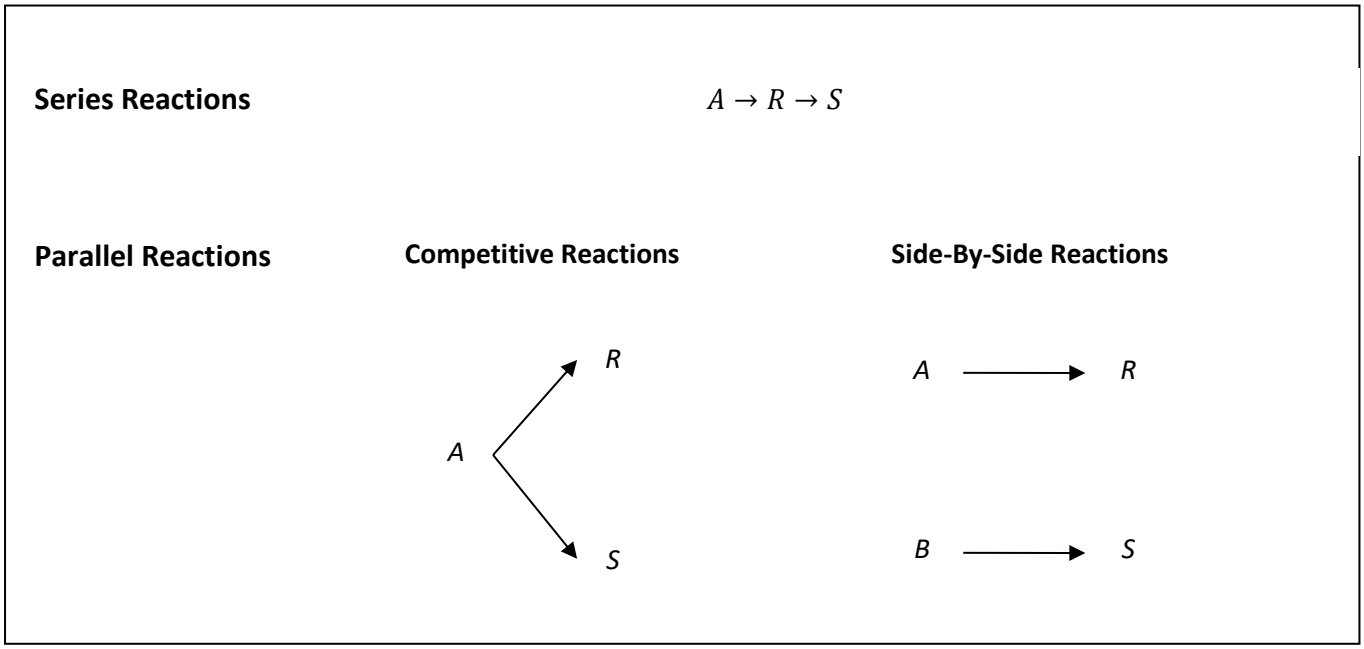


Figure 2-11. Series and Parallel Reactions

2.5.1. Elementary & Non-Elementary Reactions

An *elementary reaction* refers to a reaction step reaction that has a single transition state and no intermediates products (Nguyen and Ngo, 2020). However, *non-elementary/ complex reactions* consist of two or more *elementary reactions* that can occur in series or parallel. The *reaction mechanism* for a *complex reaction* is the set of *elementary reaction steps* that occur (Levenspiel, 1999). The concept of *elementary* or *non-elementary reactions* is illustrated in Figures 2-12 and 2-13. The clear difference between Figures 2-12 and 2-13 is the presence of one or more (n) intermediates and two or more (n+1) transition states associated with *complex reactions*. The reason why complex reactions are represented as overall reactions, with intermediates not shown, is because the number of intermediates that form are small and often destabilize to form products very quickly (Levenspiel, 1999).

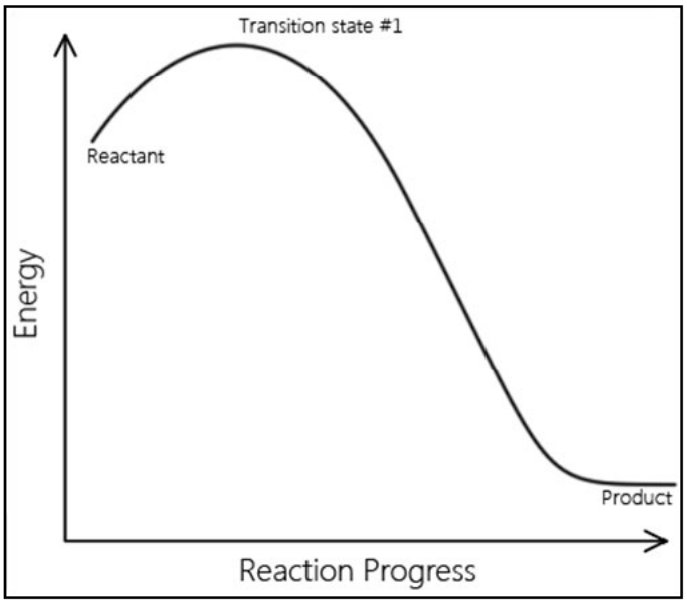


Figure 2-12. Elementary Reaction Coordinate (Nguyen and Ngo, 2020)

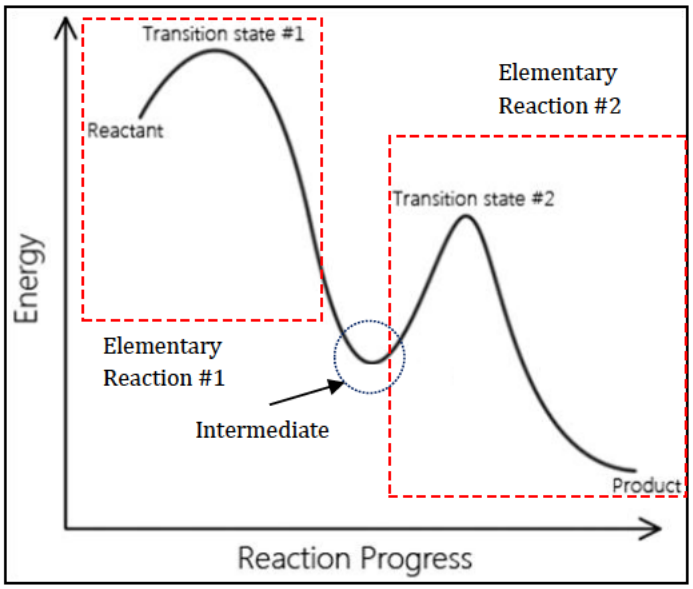


Figure 2-13. Non-elementary/ Complex Reaction Coordinate (Nguyen and Ngo, 2020)

The metathesis mechanism shown in Figure 2-4 would fall under the non-elementary classification of reactions, as it passes through multiple transition states.

2.6. Reaction Molecularity

The *molecularity* of an *elementary reaction* refers to the number of reactant molecules required to form a single molecule of the product/s (Levenspiel, 1999). This is important when it comes to understanding the mechanism of a reaction at the *molecular level*. Table 2-1 provides a technical representation of the reaction *molecularity* for an *elementary reaction*.

2.6.1. Unimolecular Reactions

Unimolecular reactions generally involve a single molecule rearranging itself to form a product. Examples of such reactions are *isomerization*, *ring-opening* and *thermal decomposition* (Nguyen and Ngo, 2020). However, the *Lindeman mechanism*, which is described in Section 2.7.1, illustrated that some *unimolecular* reactions undergo an activation step that allows for *bimolecularity* in the reaction (Nguyen and Ngo, 2020). This is particularly important when modeling the rate constant in gas phase reactions, as reactant molecule collisions must occur as required by the *collision theory*.

2.6.2. Bimolecular Reactions

Bimolecular reactions involve the collision of two reactant molecules, and commonly occur in organic reactions (Nguyen and Ngo, 2020).

2.6.3. Termolecular Reactions

Termolecular reactions involve the collision of three reactant molecules; however, it is not very common since more reactant molecules must be in the correct orientation, at the same location, and possess the required amount of energy (Nguyen and Ngo, 2020).

Table 2-1. Types of Elementary Reactions (Nguyen and Ngo, 2020)

Molecularity	**Possible Elementary Steps**	**Typical Rate Law (Elementary)**
Unimolecular	$A \rightarrow Products$	$rate = kC_A$
Bimolecular	$A + A \rightarrow Products$	$rate = kC_A{}^2$
	$A + B \rightarrow Products$	$rate = kC_AC_B$
Termolecular	$A + A + A \rightarrow Products$	$rate = kC_A{}^3$
	$A + A + B \rightarrow Products$	$rate = kC_A{}^2C_B$
	$A + B + C \rightarrow Products$	$rate = kC_AC_BC_C$

2.7. Non-elementary Reactions & Kinetic Models

Non-elementary reactions consist of more than one step; therefore, a kinetic model must capture the effect of each step. As discussed in Section 2.5.1, *non-elementary* reactions are composed of several *elementary* reaction steps, with reactive *intermediates*. The *intermediates* are formed in very small quantities and exist for very short periods of time. Therefore, only the initial reactants and final products are represented in the overall reaction (Levenspiel, 1999).

As illustrated in Section 2.1.1, the *metathesis mechanism* for a single reaction is composed two *elementary reactions*, which occur in *series* with respect to the formed *alkylidenes*. Therefore, the overall Metathesis mechanism is a non-elementary reaction.

The types of *intermediates* that can form are dependent on the chemistry of reacting species. According to Levenspiel (1999), these *intermediates* can be classed as,

1. Free Radicals - Atoms or molecules that contain one or more unpaired electrons.
2. Ions and Polar Substances – Electrically charged atoms or molecules.
3. Molecules – Those that are highly reactive, which results in a very small mean lifetime during a reaction.
4. Activated/ Transition state complexes – Molecules formed by strained or unstable bonds that can either decompose to form products or return to its original state.

The reaction mechanisms that involve these *intermediate* types are classified as either *chain* or *non-chain* reactions. The main difference between these mechanisms is that *chain* reactions involve an extra propagation step, which catalyzes the reactant to product conversion, before the *common process* of *intermediate* decomposition to the product state occurs (Levenspiel, 1999).

The *activated complex* formation is the most applicable to this work, and these generally follow a *non-chain reaction scheme* (Levenspiel, 1999).

2.7.1. Activated Complex Intermediates: Non-chain Mechanisms

Consider the hypothetical reaction,

$$A \rightarrow P + Q \ldots (1)$$

If this reaction is known to form an *activated complex intermediate* (A^*), then the following type of mechanism is postulated (Levenspiel, 1999),

- $A + A \rightarrow A^* + A$ - Energized/ Unstable molecule (*Activated complex*) formation.
- $A^* + A \rightarrow A + A$ - Return to stable form through collision.
- $A^* \rightarrow P + Q$ - Decomposition of *Activated complex* into products.

The illustration presented is shown for a *unimolecular reaction* and is also known as the *Lindemann mechanism* (Levenspiel, 1999). These types of mechanisms can easily be extended to *bimolecular* and *termolecular*

reactions if they are known to form *activated complex intermediates*. The activated complex exists for only a short period of time; however, its formation often controls the chemical reaction pathway. Therefore, it must be adequately accounted for in kinetic modeling approaches.

2.8. Packed Bed Reactor Models

Consider the overall reversible *metathesis* reaction,

$$A + B \leftrightarrow C + D \ldots (2)$$

The metathesis process involves reactions occurring by means of a *catalytic mechanism*. The overall *mass balance* for a reacting chemical system is as follows,

$$In - Out + Generation = Accumulation \ldots (3)$$

For the reactant chemical species A, with 1:1 stoichiometry, the finite element *mole balance* becomes,

$$F_{Ao} - F_A + G_A = \frac{\Delta N_A}{\Delta t} \ldots (4)$$

$$F_{Ao} - F_A + r_A \Delta W = \frac{\Delta N_A}{\Delta t} \ldots (5)$$

$$F_{Ao} - F_A + r_A \Delta W = \frac{\Delta N_A}{\Delta t} \ldots (6)$$

Assuming that the reacting system is at steady state $\left(\frac{\Delta N_A}{\Delta t} = 0\right)$, occurring in an *isothermal packed bed reactor,* with no pressure drop, then the species balance for A becomes,

$$F_{Ao} - F_A + r_A \Delta W = 0 \ \ldots (7)$$

$$\lim_{\Delta W \to 0} \frac{\Delta F_A}{\Delta W} = \frac{dF_A}{dW} = r_A \ \ldots (8)$$

Where, r_A is the overall/ net rate of consumption and generation of species A, W is the catalyst mass, and F_A is the molar flowrate of species A.

For the other reacting species, a similar process can be followed to yield (Levenspiel, 1999),

$$\frac{dF_B}{dW} = r_B \ \ldots (9)$$

$$\frac{dF_C}{dW} = r_C \ \ldots (10)$$

$$\frac{dF_D}{dW} = r_D \ \ldots (11)$$

Where, r_A, r_B, r_C & r_D can have negative or positive values, depending on the interactions in the reacting system. If multiple reactions that consume or generate species A, B, C or D occur, then these effects can be captured as part of the overall reaction rate, as follows,

$$r_A = \sum_i r_{A,formed,i} - \sum_j r_{A,consumed,j} \quad \ldots (12)$$

Where, i refers to reactions in which A is formed, and j refers to reactions in which A is consumed. This generalised expression can be extended to all chemical species that form part of the reacting system. Considering that a reaction system will be modelled, with the focus being on isolating the effects of the metathesis reactions, this understanding is a key aspect of this work. Therefore, reaction rate expressions are also reviewed in Section 2.9 to provide more context into the use of kinetic parameters.

2.9. Reaction Rate Expressions

For a hypothetical *homogenous reversible elementary reaction* of the form,

$$aA + bB \underset{k_2}{\overset{k_1}{\rightleftarrows}} rR + sS \quad \ldots (13)$$

The *overall rate* of reaction, in terms of *species-specific reaction rates*, is described as follows,

$$rate = \frac{-r_A}{a} = \frac{-r_B}{b} = \frac{r_R}{r} = \frac{r_S}{s} \quad \ldots (14)$$

The rate of a reaction is influenced by the operating temperature and concentration of species in the reaction mixture (Levenspiel, 1999). Therefore, the *overall rate law* takes the following general form for *elementary reactions*,

$$rate = k_1 C_A^a C_B^b - k_2 C_R^r C_S^s \quad \ldots (15)$$

Where, the temperature dependency is captured by the *rate constant (k)* and the concentration dependency is captured by the *species concentrations (Ci)*. The general form of a *non-elementary* reaction is different; however, the same temperature and species concentration dependencies exist. *Non-elementary* rate forms are discussed in Section 2.7. Reaction rates can be presented on different bases, i.e., area, volume, or mass of catalyst; this is for convenience in use when different types of data are available/ required.

2.9.1. Reaction Order and Rate Constant

Assuming that a hypothetical *irreversible* reaction progresses as follows,

$$A + B + C \rightarrow Products \quad \ldots (16)$$

Its *overall reaction rate* will be defined using a rate law of the form,

$$rate = kC_A^a C_B^b C_C^c \quad \ldots(17)$$

Where, *k* is the rate constant, and *a*, *b* and *c* represent the order of the reaction with respect to components *A*, *B* and *C*, respectively. The values of *a*, *b* and *c* do not need to match the stoichiometry, as they are derived from an empirically determined rate law. The overall order of the reaction (*n*) is given by,

$$n = a + b + c \quad \ldots(18)$$

It is important to note that the *reaction order* is derived from an empirically determined rate expression, which implies that it can take fractional values, whilst the *molecularity* must be an integer value as it is derived from the *reaction mechanism* of an *elementary reaction* (Levenspiel, 1999). As indicated in Table 2-2, the *order* of *elementary reactions* agrees with the stoichiometry of the *elementary step* (Levenspiel, 1999).

The *rate constant (k)* for a *homogenous* chemical reaction defines the rate at which a reaction progresses. It is strongly related to the temperature at which a reaction occurs. The *rate constant* units vary depending on the *order* of a reaction and the desired units of a *reaction rate*.

Table 2-2. Rate Constant units for different Reaction Orders and Reaction Rate Units

Reaction Rate units	**Rate Constant (k) units for different Reaction Orders**		
	1st Order	**2nd Order**	**3rd Order**
Catalyst Area based $\left(\frac{mol}{m^2.sec}\right)$	$\frac{m}{sec}$	$\frac{m^4}{mol.sec}$	$\frac{m^7}{mol^2.sec}$
Catalyst Volume based $\left(\frac{mol}{m^3.sec}\right)$	$\frac{1}{sec}$	$\frac{m^3}{mol.sec}$	$\frac{m^6}{mol^2.sec}$
Catalyst Mass based $\left(\frac{mol}{kg.sec}\right)$	$\frac{m^3}{kg.sec}$	$\frac{m^6}{mol.kg.sec}$	$\frac{m^9}{mol^2.kg.sec}$

Note: Concentrations are assumed to be in units of $\frac{mol}{m^3}$

Numerous theories have been developed for the modeling of the *rate constant*, due to its importance in the reactor design phase. These theories, together with their applicability conditions, are discussed in Section 2.10 and 2.11.

2.10. Collision Theory

Collision theory was developed using the *kinetic molecular theory* of gases as a basis; therefore, it is most applicable to gas phase behaviour. According to Fleming (2020), *collision theory* models rate constants in terms of the following,

- The frequency of molecular collisions.
- The energy molecules possess upon collision, i.e., it must be approximately greater than the *activation energy* of the reaction.
- The relative orientation of the molecules involved in a collision.

These factors dictate the rate at which the reaction will occur, as a reduction in collision frequency, insufficient energy upon collision, or improper orientation will result in a reduction in the reaction rate. The energy required is used to break existing reactant bonds and form the new product molecule bonds. *Collision theory* fails to explain reactions that do not require collisions in order to occur, i.e., *unimolecular reactions*; however, according to the *Lindeman mechanism*, from Section 2.7.1, most of these reaction types behave *bimolecularly* in the activation step which implies that collisions are indeed required (Fleming, 2020).

2.11. Transition State Theory

Transition state theory (TST) aims to provide a more accurate method to predict the rate constant, compared to the *collision theory* (Dao, Latif and Zhao, 2020). TST considers molecules to have a defined structure, unlike *collision theory* which assumes that reactant molecules are hard spheres that react upon collision, if certain prerequisite conditions are satisfied (Dao, Latif and Zhao, 2020). This assumption of the *collision theory* lacks a probabilistic element, i.e., the chance of a reaction occurring once a suitable collision has occurred.

According to the TST, there exists a state known as the *transition state*, between the reactant and product states. At the *transition state*, the reactants are combined to form the *activated complex* (Dao, Latif and Zhao, 2020). In addition, TST describes the rate of a chemical reaction, by assuming *quasi-equilibrium* between the reactants and the *activated complex* (Ptáček *et al.*, 2020). Even though this equilibrium is different from the traditional type, it can still be treated using the same thermodynamic methods (Ptáček *et al.*, 2020).

The occurrence of a reaction, i.e., product formation, is dependent on the way in which the *activated complex* separates. According to Dao, Latif and Zhao (2020), the TST provides the following factors that determine whether a reaction will occur,

- The *activated complex* concentration.
- The rate of *activated complex* separation.
- The way the *activated complex* separates. It could break apart to form the products, reform the reactants or form a new chemical complex.

For the *activated complex* to form, a certain amount of energy (*activation energy*) is required to overcome the existing reactant bonds and create the *activated complex* (unstable), which will ideally break apart to form the products. The desired activated complex will only form when reactant molecules are properly oriented and collisions are energetic enough to overcome the *activation energy* (Ptáček *et al.*, 2020). Therefore, the TST builds on the concept of collisions and energy transfer from the *collision theory* but introduces a probabilistic element that accounts for the chance of product molecule formation from the transition state complex (Dao, Latif and Zhao, 2020). A typical potential energy diagram is provided in Figure 2-14.

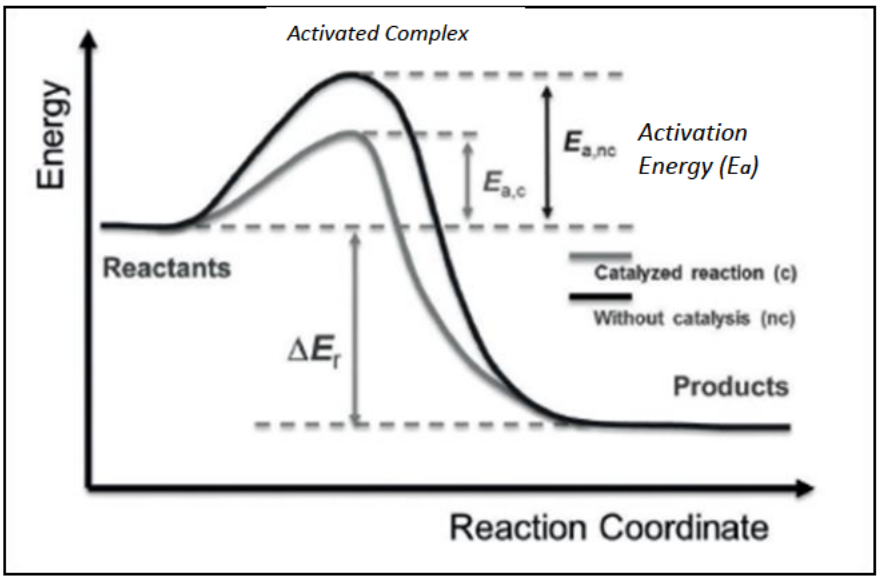


Figure 2-14. Typical Reaction Potential Energy Diagram (Ptáček *et al.*, 2020)

2.11.1. Limitations of the TST

The transition state theory is limited to the one step mechanisms illustrated in Figure 2-12 and Figure 2-14. Typical cases for which the TST fails includes (Ptáček *et al.*, 2020),

1. If the TST is applied to an elementary step of a multistep reaction.
2. When the intermediates exist for a very short period, i.e., the Boltzmann energy distribution is not reached before the process moves to the next step.
3. When the process operates at very high temperatures (*complex molecular motion*) or very low temperatures (*quantum tunneling*).

Quantum tunneling refers to reactant molecules transformation into the *activated complex*, despite having insufficient energy upon molecular collision (Ptáček *et al.*, 2020).

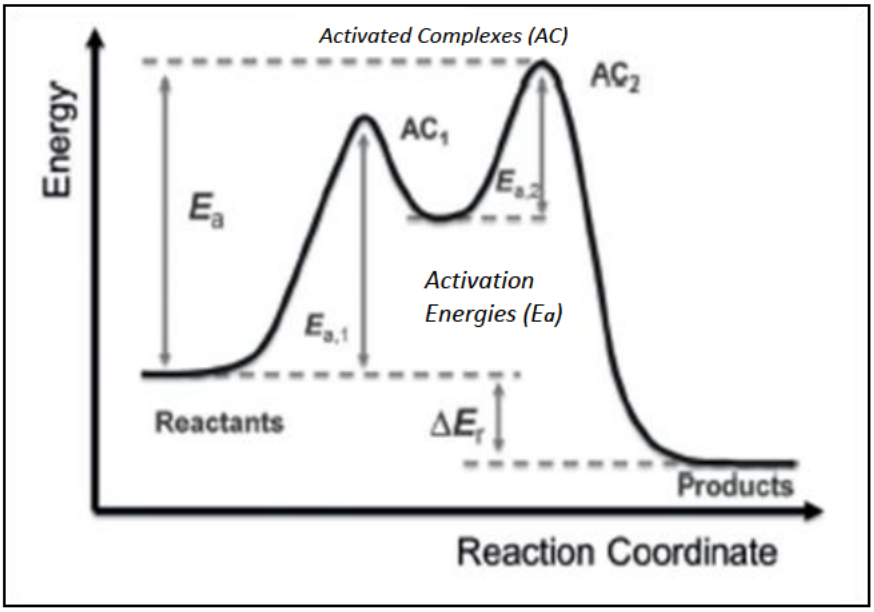


Figure 2-15. Multistep Reaction Potential Energy Diagram

For mechanisms such as those illustrated in Figure 2-15, where the reaction must pass through two steps/ activated complexes, *point 1* of the TST limitations occurs. This is typical in the metathesis mechanism, in Section 2.1, as different *alkylidenes* are required to complete one metathesis cycle; therefore, standard methods based on the TST cannot be readily applied.

2.11.2. The Eyring Equation

The Eyring equation defines the temperature dependence of reaction rate constants, in a manner similar to the Arrhenius equation (Louie, 2020). However, it accounts for the energy barrier associated with the formation of the transition state, which increases accuracy and accounts for microscopic factors involved in a reaction (Louie, 2020). The Eyring equation is applicable to gas, liquid, and mixed phase kinetics, whereas the Arrhenius equation is only applicable to gas phase kinetics (Louie, 2020). The Eyring equation is described as follows,

$$k = \frac{k_B T}{h} e^{-\frac{\Delta G^{+}}{RT}}$$

$$k = ve^{-\frac{\Delta G^{+}}{RT}}$$

Where, k is the *overall reaction rate constant*, k_B is the *Boltzmann constant* with a value of $1.381 \times 10^{-23} J/K$, T is the Temperature in Kelvin, h is *Planck's constant* with a value of $6.626 \times 10^{-34} J.s$, R is the *Universal gas constant* with a value of $\frac{8.314J}{mol.K}$, v is the vibrational frequency of the transition state, and ΔG^{+} is the *Gibbs energy of Activation* which can be visualized by the activation energy illustrated in Figure 2.14.

Another useful representation of this equation is derived from basic thermodynamic principles. The *Gibbs energy of Activation* is thermodynamically defined as,

$$\Delta G^{+} = \Delta H^{+} - T\Delta S^{+}$$

Where, ΔH^{+} is the *Enthalpy of Activation* [J/mol] and ΔS^{+} is the *Entropy of Activation* [J/mol/K]. Using this relation, the *Eyring equation* can be re-expressed as follows (Louie, 2020),

$$k = \frac{k_B T}{h} e^{-\frac{\Delta H^{+}}{RT}} e^{\frac{\Delta S^{+}}{R}}$$

To use the Eyring equation, the *Gibbs energy of Activation* (ΔG^{+}) must be evaluated. The *Gibbs energy of Activation* (ΔG^{+}) is related to the *Thermodynamic equilibrium constant of Activation* (K^{+}) by the following relation,

$$\Delta G^{+} = -RT\, lnK^{+}$$

This provides a more quantifiable method to calculate the *Gibbs energy of Activation* (ΔG^{+}), using a transition state approach.

The Thermodynamic equilibrium constant of Activation (K^{+})

Consider the 1st *reversible elementary step* of a *generalised metathesis reaction*,

$$M_1 + R_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TS_1^{+} \underset{k_4}{\overset{+\ k_3}{\rightleftharpoons}} M_2 + P_1$$

The *Elementary step equilibrium constant* (K_{eq}) is given by,

$$K_{eq} = \frac{[M_2][P_1]}{[M_1][R_1]}$$

Whereas the *Thermodynamic equilibrium constant of Activation* (K^{+}) is given by,

$$K_A^{+} = \frac{[TS_1^{+}]}{[M_1][R_1]} = \frac{k_1}{k_2}$$

$$K_B^{+} = \frac{[M_2][P_1]}{[TS_1^{+}]} = \frac{k_3}{k_4}$$

Consider the 2nd reversible *elementary step* of the *generalised metathesis reaction*,

$$M_2 + R_2 \underset{k_6}{\overset{k_5}{\rightleftharpoons}} TS_2^+ \underset{k_8}{\overset{k_7}{\rightleftharpoons}} M_1 + P_2$$

The *Elementary step equilibrium constant* (K_{eq}) is given by,

$$K_{eq} = \frac{[M_1][P_2]}{[M_2][R_2]}$$

Whereas the *Thermodynamic equilibrium constant of Activation* ($\overset{+}{K}$) is given by,

$$\overset{+}{K_C} = \frac{[TS_2^+]}{[M_2][R_2]} = \frac{k_5}{k_6}$$

$$\overset{+}{K_D} = \frac{[M_1][P_2]}{[TS_2^+]} = \frac{k_7}{k_8}$$

Therefore, *Thermodynamic equilibrium constant of Activation* ($\overset{+}{K}$) defines the equilibrium that occurs between the *reacting species* and the *Transition state*.

Transition State Dissociation

The dissociation of the transition state complex it a probabilistic event. There are three possibilities when considering the dissociation of TS, these are (Perez-Benito, 2017),

1. The TS dissociates to form the products (probability P_f).
2. The TS dissociates to back into the reactants (probability P_r).
3. The TS does not dissociate due to insufficient vibrational energy (probability P_o).

The probabilities represent the likelihood of TS dissociation occurring in that way. A useful property of the probabilistic representation is,

$$P_f + P_r + P_o = 1$$

Usually, P_o is neglected so the relation becomes,

$$P_f + P_r = 1$$

For the transition state TS_1, the *dissociation rate constants* can be represented using the probabilities as follows,

$$k_2 = P_{r,TS_1} v$$

$$k_3 = P_{f,TS_1} v$$

For the transition state TS_2, the *dissociation rate constants* can be represented using the probabilities as follows,

$$k_6 = P_{r,TS_2} v$$

$$k_7 = P_{f,TS_2} v$$

Where, v is the *frequency of vibration* (or reciprocal lifetime) of the transition state complex, and is a temperature dependent parameter defined as,

$$v = \frac{k_B T}{h}$$

Where, k_B is the *Boltzmann constant* with a value of $1.381 \times 10^{-23} J/K$, T is the Temperature in Kelvin, h is *Planck's constant* with a value of $6.626 \times 10^{-34} J.s$, R is the *Universal gas constant* with a value of $\frac{8.314 J}{mol.K}$. The transition state approach is an inherently statistical approach to the modeling process and introduces a significant degree of complexity into any modeling approach.

2.12. Testing Kinetic Models

Many factors make it difficult to find the correct mechanism of a reaction. Firstly, reactions may proceed by a single mechanism, i.e., it involves only one group of *intermediate*, or a combination of different mechanisms, i.e., it involves more than one class of *intermediate* (Levenspiel, 1999). Secondly, many mechanisms can be consistent with the available/ experimental kinetic data. The true mechanism can be determined through an understanding of the chemistry of all species involved in the reaction (Levenspiel, 1999). The mechanism of a reaction involves a sequence of elementary reactions, in which there hypothetically exists one of the following two types of intermediates (Levenspiel, 1999),

2.12.1. Type 1 Intermediates

Type 1 intermediates (X) are usually present in small concentrations, such that the rate of change of its concentration in the mixture is negligible, i.e., zero (Levenspiel, 1999). This known as the *steady-state approximation*, and is represented as,

$$[X] \text{ is small and } \frac{d[X]}{dt} \cong 0$$

Where, [X] represents the concentration of intermediate X.

2.12.2. Type 2 Intermediates

Type 2 intermediates (X) form when a *homogenous free catalyst (C)* is 'consumed' in a reaction. The homogenous catalyst is present in an *initial concentration* C_O. Therefore, the following relation must hold,

$$[C_O] = [C] + [X]$$

Where, [X] represents the concentration of intermediate X.
According to Levenspiel (1999), the following assumption is also applicable,

$$\frac{d[X]}{dt} = 0$$

Or the intermediate (X) can be assumed to be in equilibrium with its reactants; therefore,

$$\binom{reactant}{A} + \binom{catalyst}{C} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \binom{intermediate}{X} \Rightarrow K_e = \frac{k_1}{k_2} = \frac{[X]}{[A][C]}$$

Where, k_1 and k_2 are the rate constants for the forward and reverse reactions, $[i]$ represents the concentration of the *ith species*, and K_e is the equilibrium constant.

The type of intermediate is different from the class of intermediate. In general, one group of intermediates can be only a specific class of intermediate.

The metathesis mechanism is well understood from literature and is illustrated in Figure 2-4. From Figure 2-4, it fits in with both the Type 1 and Type 2 intermediates – Firstly, Type 2 intermediates are formed during the induction period, then the type 1 mechanism occurs with conventional metathesis processes.

2.13. Arrhenius Law

Arrhenius' law describes the temperature dependency of the rate constant, and is described as follows,

$$k = A_O e^{\frac{-E_a}{RT}}$$

The equation above is known as the *Arrhenius equation*, in which, k is the reaction rate constant, A_O is the frequency/ pre-exponential factor, R is the universal gas constant, T is the temperature in *Kelvin* and E_a is the activation energy of the reaction. *Arrhenius' law* is known to provide a good approximation to the actual temperature dependency of the rate constant, in accordance with both *collision theory* and *transition state theory* (Levenspiel, 1999).

2.13.1. Activation Energy

The activation energy represents the minimum energy a reactant molecule must possess to break existing bonds and form the products (Fogler, 2006). Reactions can be either *exothermic* or *endothermic*; however, the activation energy will remain positive and approximately constant during a reaction. The *exothermic* or *endothermic* nature of a reaction depends on the *heat of reaction* (ΔH_R). Some useful facts about the *potential energy diagram* in Figure 2-14, are listed below,

1. Positive *heats of reaction* (ΔH_R) imply that energy is required for a reaction to occur, i.e., endothermic, and negative *heats of reaction* (ΔH_R) imply that energy is released during a reaction, i.e., exothermic.
2. The higher the activation energy, the greater the energy that will be required to make the reaction proceed from reactants to products – It also implies that the reaction becomes more *temperature-sensitive* (Levenspiel, 1999). This implies that a *larger temperature rise* would be required to bring about an equivalent increase in a specific reactions rate, compared to another reaction with lower activation energy.
3. Reactions are generally more *temperature-sensitive* at lower operational temperatures (Levenspiel, 1999).

2.13.2. Pre-Exponential/ Frequency Factor

This represents a statistical measure of the number of times reactant molecules are likely to collide with the orientation necessary to form the product molecule. The pre-exponential factor can be empirically determined, using the *Arrhenius method*, or it can be described using other methodologies, such as *transition state theory* or *collision theory*.

2.13.3. The Role of Arrhenius' Law in Kinetic Modeling

For the purpose of kinetic modeling, both A_O and E_a are required and can be obtained using a modified form of the *Arrhenius equation*, assuming some rate constant and temperature data combinations are known. The equation is modified using the following transformations,

Taking the natural logarithm on both sides of the *Arrhenius equation* yields,

$$ln(k) = \left(-\frac{E_a}{R}\right)\left(\frac{1}{T}\right) + ln(A_O)$$

This equation has a linear form, with *slope* $-\frac{E_a}{R}$, *y-intercept* $ln(A_O)$, independent variable $\frac{1}{T}$, and dependent variable $ln(k)$. The graphical representation of this relationship is provided in Figure 2-16.

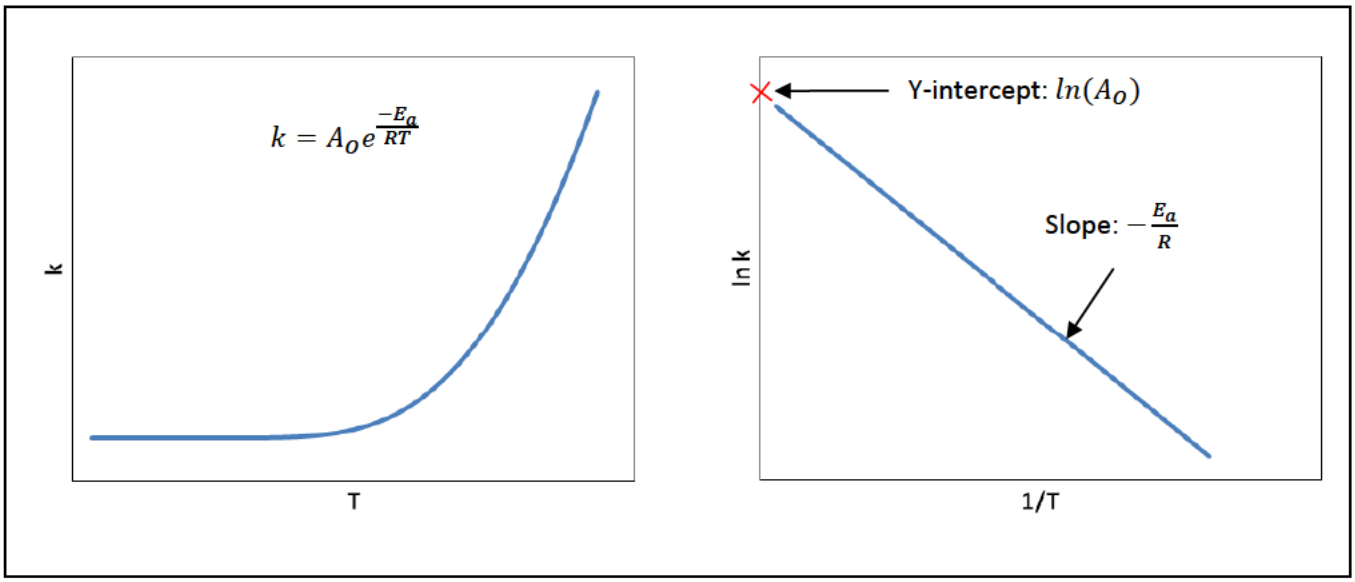


Figure 2-16. The Arrhenius Transformation

2.14. Controlling Mechanisms

Reactions can reach the same product via different mechanisms, i.e., *parallel mechanisms*, or *sequential (series)* mechanisms. From *Section 2.5*, reaction mechanisms indicate the path taken from reactant to product in terms of the elementary reaction steps involved – Series or parallel mechanisms are different from series or parallel reactions. A reaction mechanism is effectively a group of series or parallel reactions.

In general, a change in the observed *activation energy* indicates a shift in the controlling mechanism of a reaction (Levenspiel, 1999). Each mechanism usually has its own *controlling step*, with an associated activation energy; therefore, different Arrhenius plot slopes are expected when a reaction consists of multiple mechanisms. The *controlling step* refers to the *slowest elementary step* in a particular mechanism and is indicated by the *highest activation energy*.

Figure 2-17 illustrates the type of *activation energy* shift, i.e., high-to-low (concave) or low-to-high (convex), for the overall conversion of species A to R. This shift relates to whether the possible mechanisms of that reaction occur in series or parallel.

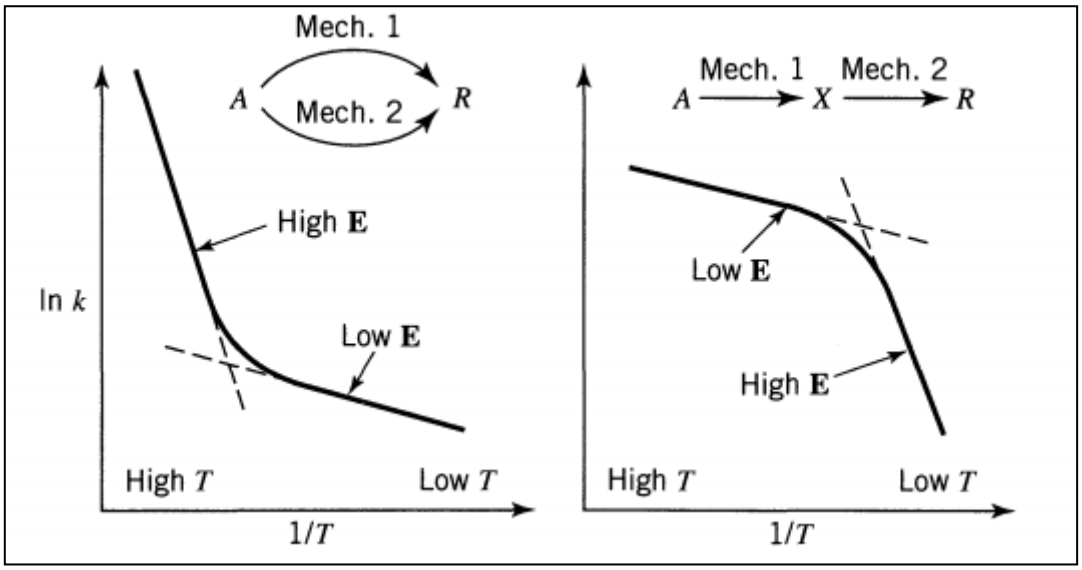


Figure 2-17. Representation of Typical Controlling Mechanism Shifts (Levenspiel, 1999)

Parallel mechanisms are not restricted to occur in a specific order; therefore, at low temperatures the *lower activation energy* mechanism will be controlling, as it is the only mechanism that likely has sufficient energy to occur. At higher temperatures, where both mechanisms have sufficient energy to occur, the *higher activation energy* mechanism will be controlling as it will contain the slowest step. This interpretation is based on the fact that the *activation energy* dictates the amount of energy required for a reaction to occur; therefore, only a mechanism with sufficient energy to occur will take place.

Series mechanisms occur in a specific order; therefore, there is no guarantee that a higher temperature will cause the *higher activation energy* mechanism to become controlling, as the second mechanism is dependent on the occurrence of the first, as shown in Figure 2-17. This will translate to a lower-than-expected rate at high temperatures (Truhlar and Kohen, 2001).

Tolman's interpretation of the *activation energy* is often used to understand these activation energy shifts, i.e., it indicates the reason behind the occurrence of *convex* or *concave Arrhenius plots*. However, *Tolman's* interpretation is strictly applicable to *elementary reactions*, which exist within a controlling mechanism. *Tolman's* interpretation states,

$$E_a = \bar{\bar{E}} - \bar{E}$$

Where, E_a is the activation energy, $\bar{\bar{E}}$ denotes the average energy of all molecules that react and $\bar{E}$ denotes the average energy of all possible reactants.

In the case of *concave Arrhenius plots*, for an increase in temperature $\bar{\bar{E}}$ increases more rapidly than $\bar{E}$, which is attributed to a rate constant that increases with temperature (Truhlar and Kohen, 2001). However, in the case of *convex Arrhenius plots*, for an increase in temperature $\bar{\bar{E}}$ decreases or increases less rapidly than $\bar{E}$, which is attributed to a rate constant that decreases with temperature (Truhlar and Kohen, 2001). Truhlar and Kohen (2001) further explained that *convex Arrhenius plots* arise when molecules traverse a wider region of phase space that contains a larger number of non-reactive conformers (stereoisomer) and are less likely in non-biological reactions.

Therefore, there are two factors that explain the non-linearity that can be observed in Arrhenius plots, and that is,

1. When a combination of several series and/or parallel mechanisms occur in a particular reaction.
2. When the activation energy of the controlling step of an *elementary*, single mechanism, reaction varies with temperature as explained by *Tolman's* interpretation.

It is also possible for a combination of these factors to occur. These factors may also be used to determine the level at which a reaction is being modeled – Elementary reactions will result in linear Arrhenius plots, and non-elementary reactions will result in non-linear Arrhenius plots.

2.15. Optimization Algorithm: Particle Swarm Optimization

The *particle swarm optimisation (PSO) algorithm* is a type of population-based metaheuristic optimisation technique – It uses swarm intelligence to search for an input solution that meets the required objective (Mirjalili *et al.*, 2020). These operate by sampling the search space of a problem that is too large to be searched completely. A typical computational set-up of the algorithm is shown in Figure 2-18. The problem with metaheuristic techniques is that the solutions determined can be partially true, or even inaccurate (Mirjalili *et al.*, 2020).

Input: Population Size (N_p), Number of Iterations (*num_iter*), Constants c_1, c_2 and w

Output: Global best solution (P_{global})

Process:

Generate initial population with Np individuals P_i ($i = 1, 2, 3, \ldots, N_p$)

Set iteration counter t = 0

Compute the fitness of each Pi, assign P_{global} as overall individual & P_{local} as population with best fitness

While (t < num_iter)

For each individual (P_i),

Calculate velocity vector (v_i), with c_1, c_2, and w_i

Apply velocity vector (v_i) to update each individual (P_i)

Compute the fitness (f_i) of each individual (P_i)

Update $P_{local,\,i}$ & $f_{local,\,i}$ if $f_i < f_{local,\,i}$

Update P_{global} & f_{global} if $f_i < f_{global}$

End

Increment t by 1

End While

Figure 2-18. Classical PSO Algorithm (Mirjalili *et al.*, 2020)

The power of the PSO algorithm lies in its ability to incorporate a social (global) component, in addition to the individual (local) components present in optimisation algorithms. In a typical set-up, the population would consist of individuals (N_p), which in turn consist of the parameters (n dimensions) that make up that individual. The values of these parameters are analogous to the position of that individual in an n dimensional space.

The PSO technique introduces a velocity vector that allows the parameters of an individual to update – This allows the individuals to move through the n dimensional space, with influence from factors that involve the individual best (local component), and global best (social component). The velocity vector can be calculated as shown in Figure 2-19.

The following equations can be used to determine the future velocity vector ($V_i|_2$) to be used to update the ith individuals in a population (P_i),

$$V_i|_2 = w.V_i|_1 + c_1.r_1\left(P_{local,i}|_1 - P_i|_1\right) + c_2.r_2\left(P_{global}|_1 - P_i|_1\right)$$

$$P_i|_2 = P_i|_1 + V_i|_2$$

Where, r_1 & r_2 are random numbers between 0 and 1, w is the inertial weight, c_1 is the individual coefficient, and c_2 is the social coefficient.

Figure 2-19. PSO n-parameter velocity vector calculation process (Mirjalili *et al.*, 2020)

Figure 2-19 shows that the velocity vector consists of multiple components, these are,

- The tendency toward current velocity → $w.V_i|_1$
- The individual intelligence → $c_1.r_1\left(P_{local,i}|_1 - P_i|_1\right)$
- The social intelligence → $c_2.r_2\left(P_{global}|_1 - P_i|_1\right)$

The impact of each of these components can be controlled by using the relevant constants (w, c_1 & c_2); therefore, it is highly important that these values be adequately selected. A degree of diversity is introduced by using the random numbers r_1 & r_2; however, these only minimally influence the rate of change related to each component. Mirjalili *et al.* (2020), investigated the impact of each of these constants in a generalised context – It was found that the *inertial weight (w)* should be set to a value that linearly decreases from 0.9 to 0.4 proportional to the number of iterations, and the parameters (c_1 & c_2) should be set to a value between 1 & 2 to ensure fast convergence and a more global search range. It should be noted that these are recommendations and can be modified based on the observed optimisation behaviour.

Further findings from the work of Mirjalili *et al.* (2020), was that the value of the *inertial weight* should never be set to zero as this would cause low *exploitation* (local search); instead, a dynamically changing value is preferred to balance high and low exploitation. The value of the *cognitive component (c_1)* should never be set above the value of 3, as this would cause particles to gravitate toward their personal best parameters. The *social component (c_2)* achieves a reasonable balance between local and global search when it is set to a value of 2.

2.16. Optimization Algorithm: Genetic Algorithm

The *genetic algorithm (GA)* is a type of population-based metaheuristic optimization technique – It effectively creates a black box model around the system and uses evolutionary principles to search for an input solution that meets the required objective (Katoch, Chauhan, and Kumar, 2020). The classical GA is set-up as shown in Figure 2-20.

Input: Population Size (N_p), Max Number of Iterations (*num_iter*)

Output: Global best solution (Y_{bt})

Process:

Generate parent population of Np chromosomes Y_i (I = 1, 2, 3, …, N_p)

Set iteration counter t = 0

Compute the fitness of each Yi

While (t < num_iter)

Select pairs of chromosomes from parent population

Apply crossover operation on selected pair, with crossover probability, to generate offspring

Apply mutation on the offspring, with mutation probability, to update offspring

Replace old population, with newly generated population, using survival

Increment t by 1

End While

Figure 2-20. Classical Genetic Algorithm (Katoch, Chauhan, and Kumar, 2020)

The GA involves creating predictions using high variability and are excellent at searching a well-defined global space, however, they are heavily driven by probability and randomness (Katoch, Chauhan, and Kumar, 2020).

The GA consists of several *genetic operators*, these are shown below (Katoch, Chauhan, and Kumar, 2020),

- Random generation and encoding of a population (P);
- Selection of individuals or sets within P;
- Randomized crossover within P, to create offspring (O);
- Randomized mutation of O to give rise to an updated O;
- Survival through deterministic selection of the best individuals from the overall set of O and P.

The *encoding step* is of particular importance when using a GA, as the information needs to be put into a format that can easily be used and implemented in the simulation, but still retains the information necessary to add meaning to the system model. For system models that involve chemical/ realistic processes, *real value* encoding is preferred as these often use very low order of magnitude numbers, which are difficult to encode using other schemes – The benefit of real coded GAs is that they are robust, efficient, and accurate (Katoch, Chauhan, and Kumar, 2020).

Selection is a genetic operator that determines whether a certain participant in the population will be involved in the evolutionary process. Various selection techniques exist, and these are based on random selection of participants from the population. *Tournament selection* is a popular technique and is also introduces a degree of randomness into the process – It involves selecting participants in pairs, and comparing those based their individual fitness values to determine which would be included in the next steps of the process. This process is repeated until an updated population that consists of the same number of participants as the original population. A lower number of tournaments are known to create genetic diversity (Katoch, Chauhan, and Kumar, 2020).

The next genetic operator is *crossover* – This involves combining the selected participants to produce an offspring. The technique used to implement the crossover process is directly related to the encoding technique applied to the population – Given that real value encoding is preferred for chemical systems, a crossover technique known as *Simulated binary crossover (SBX)* provides a useful method of simulating this process. In all cases, the likelihood of crossover is controlled by the *probability of crossing over (p_c)*. This is a fractional decision boundary that determines whether crossover will occur (Katoch, Chauhan, and Kumar, 2020). The data in Figure 2-21 can be used to determine the offspring that result from this process.

The following equations can be used to determine the individual elements (p_i & q_i) of each offspring (P & Q) that results from individuals X & Y,

$$p_i = 0.5.\left((1 - beta).x_i + (1 + beta).y_i\right)$$

$$q_i = 0.5.\left((1 + beta).x_i + (1 - beta).y_i\right)$$

Where, beta is a random value that lies in the range between zero and ∞

Figure 2-21. SBX Crossover Method (Katoch, Chauhan, and Kumar, 2020)

The offspring are then exposed to the *mutation* genetic operator – The preferred type of mutation operator for real coded GAs is known as *polynomial mutation*. This ensures diversity by introducing randomness into the elements that make up a participant of the offspring (Katoch, Chauhan, and Kumar, 2020). Similar to crossover, the likelihood of mutation is controlled by the *probability of mutating (p_m)*. This is a fractional decision boundary that determines whether mutation will occur (Katoch, Chauhan, and Kumar, 2020). *Polynomial mutation* can be performed using the method illustrated in Figure 2-22. It is highly recommended that both crossover and mutation operators are used to introduce genetic diversity, as GAs are prone to premature convergence (Katoch, Chauhan, and Kumar, 2020).

The following equation can be used to determine the mutated element (p_i) of each offspring (P),

$$p_i|_2 = p_i|_1 + (ub - lb).delta_i$$

Where, $delta_i$ is a random value that lies in the range between zero and ∞, and ub & lb are the upper and lower bound respectively

Figure 2-22. Polynomial Mutation Method (Katoch, Chauhan, and Kumar, 2020)

The *survival* genetic operator is the simplest – This involves grouping all participants from the parents and offspring, with *deterministic selection* of the best participants such that the number of participants matches the size of the original parent population.

2.17. Ring Closing Metathesis

Ring closing metathesis is a useful complementary reaction that can occur, given that the required conditions are satisfied. The importance of this process is that it creates a mechanistic link between the petrochemical and pharmaceutical industries through the metathesis process.

The cyclic structures produced in RCM possess various pharmacological properties, and there have been numerous advances in the space of catalyst tuning which can be used to control the process that occurs – This is an inherently complex area of research as the position, and type of substituents present on the cyclic structure, coupled with the size of the ring, all influence the resulting properties (Jacobsen *et al.*, 2010). Therefore, the focus of this section is to provide an overview of the potential mechanism, and any modeling attempts.

As illustrated in Figure 2-23, where ADMET refers to linear diene metathesis, RCM refers to ring-closing metathesis, and ROMP refers to ring-opening metathesis polymerization, RCM involves the closing of a linear olefin (diene) to form a cyclic structure (Sinclair *et al.*, 2017). This process involves the liberation of extraneous volatile groups (ethylene gas for terminal alkenes) and can only occur for the case of diene metathesis. For reasons introduced in Section 2.1.3, ring strain becomes a driving factor in the occurrence of RCM. Therefore, RCM would be more likely to occur with longer chain alkenes, which would produce less strained rings; however, there is an optimal hydrocarbon chain length, i.e., 5 to 8 (Sinclair *et al.*, 2017). *Grubbs catalyst* (Ruthenium based) is the most selective for the RCM mechanism, as there is a high functional group tolerance associated with these catalysts.

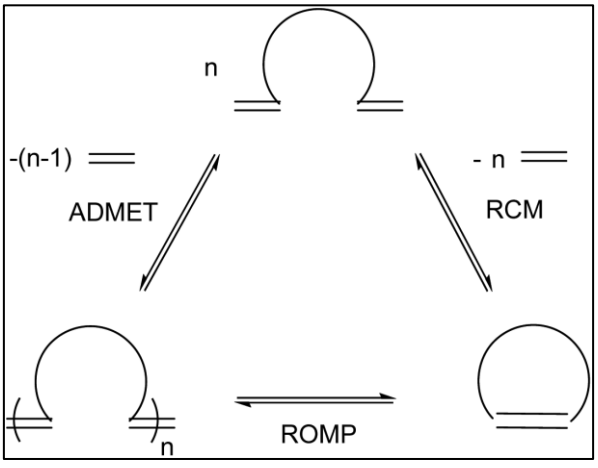


Figure 2-23. The Different Metathesis Types (Sinclair *et al.*, 2017)

Żukowska *et al.* (2013), studied the generalized olefin metathesis mechanism, which is illustrated in Figure 2-1 and 2-2, and provided a generally accepted RCM mechanism, which is illustrated in Figure 2.24. RCM proceeds quickly for moderately sized rings (5 to 7 members), but drastically slows down for larger sized rings (8 to 11 members), which will impact Ruthenium catalyst lifespan (Żukowska *et al.*, 2013).

From the mechanism in Figure 2-24, this type of metathesis cannot be directly applied to the alkene system under investigation (1-Hexene), as dienes are required to close the ring. Therefore, it will not be directly applicable to the system studied, however the methodology developed here, would be useful from a RCM modeling perspective. Given that this work is based on the formation of the transition state, and that the transition state still exists in the mechanism of Figure 2-24, the techniques to be developed in this work can be applied to the RCM mechanism with some modifications based on the observed system behaviour.

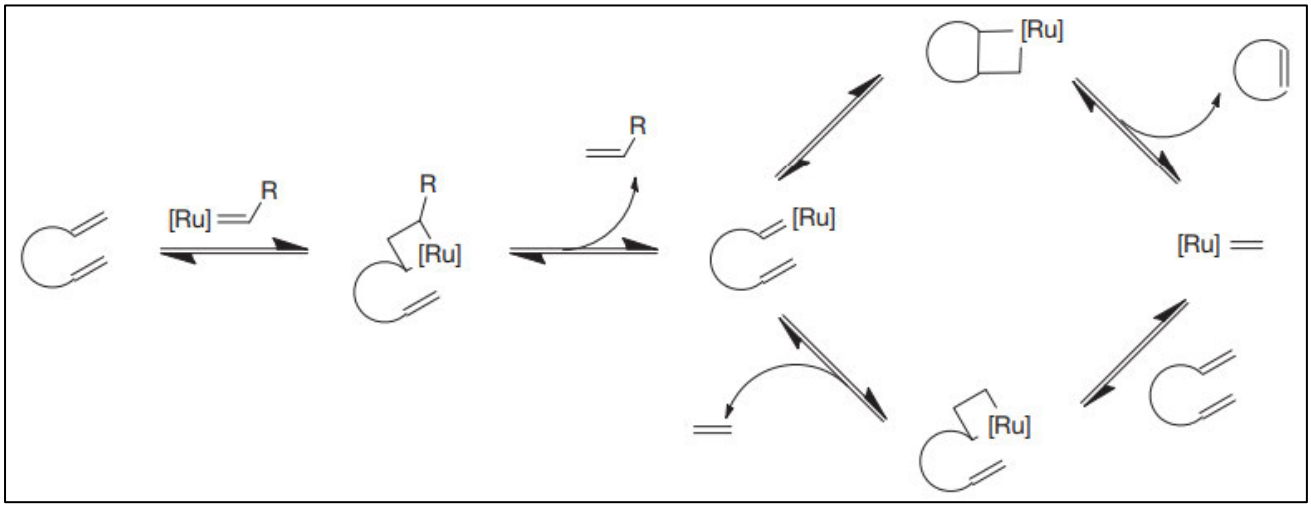


Figure 2-24. Generalized RCM Mechanism (Żukowska *et al.*, 2013)

From the work of Nelson *et al.* (2011), various methods to develop kinetic models for RCM have been tested. It was found that quantitatively described models, did not account for physical effects, and that it is important to account for induction effects in the process to ensure that adequate results are obtained. *Interchange mechanisms*, in which bonds form and break simultaneously, seem to provide a suitable way of accounting for the kinetic interactions observed; however, a more robust modeling approach is still required (Nelson *et al.*, 2011). Therefore, there exists the unique opportunity to fill this need and apply the potential methodology for the modeling and development of RCM processes.

3. The 1-Hexene Isomerisation Process

3.1. Background: Catalyst and Data Selection

Literature data from Prithipal *et al.* (2015) will be used to assess the modeling methodology developed in this work – The importance of accounting for the isomerization process is illustrated in Section 4.2, where the observation is that it vastly broadens the overall reaction network and product distributions; therefore, a 1-Hexene isomerization model is developed here for use and reference in this work.

According to Prithipal *et al.* (2015), the feed to the experimental metathesis system consisted of pure 1-Hexene gas, diluted with an inert (N_2). This creates a concentration-based driving force toward the *isomerization* of 1-Hexene. Other Alkenes, such as the Butenes or Pentenes, can also undergo *isomerization*; however, they are not present/ formed in a proportion sufficient to promote their *isomerization*. Therefore, the weak concentration-based driving force, implies that the other Alkenes do not significantly undergo *isomerization*; as such, their isomerization can be neglected under the operating conditions.

1-Hexene isomerization data, for the Tungsten trioxide (WO_3/SiO_2) catalyst, is not readily available in literature; however, Sundaramurthy (2000) obtained 1-Hexene isomerization data for Boron (B) and Aluminium (Al) substituted Zeolites (ZSMs) for different substitution ratios, i.e., 20-200. Therefore, an understanding of the catalyst properties and structure is required to find a catalyst that is most representative of the Tungsten trioxide catalyst, which will ensure that the data is most representative for use in this work. A summary of different catalyst properties is provided in Table 3-1. These come from numerous sources and are the basis for choosing a representative catalyst. The properties of the Tungsten trioxide (WO_3/SiO_2) catalyst were obtained from the work of Prithipal *et al.* (2015).

Table 3-1. Isomerisation Catalyst Comparison

Property	**WO_3/SiO_2 (8wt%)**	**Al-ZSM-5 (20)**	**B-ZSM-5 (20)**	**Al-MCM-41 (100)**	**B-MCM-41 (100)**
Pore volume [cm^3/g]	0,94	-	-	-	-
Pore size [nm]	15,65	0,5	0,5	4,1	4,1
BET surface area [m^2/g]	275	312	440	-	-
Total acidity [mmol/g]	0,2725	1,008	>>0,2779	0,2779	0,2779

The selection of a representative catalyst is governed by the catalyst structure and type of acid sites present, i.e. *Bronsted* or *Lewis* sites. *Bronsted* acid sites are much stronger compared to *Lewis* acid sites; therefore, the reactions that can occur on them differs. *Bronsted* acid sites are noted to have a preference for *isomerization* and *cracking* reactions, while *Lewis* acid sites have preference for *isomerization* and metathesis reactions (Prithipal *et al.*, 2015). According to Prithipal *et al.* (2015), WO_3/SiO_2 catalyst possesses both *Bronsted* and *Lewis* acid sites; therefore, both *isomerization* and *metathesis* are likely to occur at the experimental conditions

described in the *metathesis literature data*, of Appendix A. *Cracking* reactions generally occur with a high concentration of *saturated* longer chain hydrocarbons (> C_{12}), and high temperatures (>500°C); therefore, it would not be significant given the expected species concentrations, type, and operating conditions.

Another important point to note about the strength of acid site is that the types of isomerization reactions that can occur differ for *Lewis* and *Bronsted* sites. According to Sundaramurthy (2000), *Double bond shift* (DBS) isomerization reactions occur more easily, i.e., less energetically demanding, compared to *Skeletal rearrangement* (SR) isomerization reactions. Therefore, DBS reactions occur on the weaker *Lewis* sites, while SR reactions occur on the stronger *Bronsted* sites. Therefore, the acidity of the catalyst becomes a critical factor in the development of the isomerization kinetic model. To ensure that the type of isomerization is limited to DBS, the acidity of the representative catalyst must be as close as possible to that of the Tungsten trioxide (WO_3/SiO_2) catalyst.

3.2. Catalyst Properties

3.2.1. Catalyst 1: Al-ZSM-5 (20)

This is the most well studied catalyst of all listed in Table 3-1. It has very narrow pores (Rodríguez-González *et al.*, 2007), compared to the Tungsten catalyst, which makes mass transfer of reactant (in) and product (out) difficult. Petrik (2009) found that the BET surface area of this catalyst was $312 m^2$/g. Rodríguez-González *et al.* (2007), provided catalyst acidity data for different Al-ZSM Si/Al ratios, as illustrated in Table 3-2. Using the data in Table 3-2, Figure 3-1 was constructed, and a *power-law model* was fitted to the experimental data. For a Si/Al ratio (*X*) of 20, which corresponds to the Si/Al ratio of *Catalyst 1*, the ZSM acidity (*Y*) was estimated to be 1.008mmol/g. Al-ZSM catalysts are also both *Bronsted* and *Lewis* acidic; however, it is *3.7* times more acidic when compared to the Tungsten trioxide (WO_3/SiO_2) catalyst. This is attributed to the much higher concentration of *Bronsted* acid sites, associated with the presence of Al^{3+} (Rodríguez-González *et al.*, 2007). Therefore, SR reactions are more likely to occur with the use of this catalyst; as such, a catalyst with lower acidity is desired.

Table 3-2. Al-ZSM-5 Acidity data (Rodríguez-González *et al.*, 2007)

Si/Al Ratio	**ZSM Acidity [mmol/g]**
30	0,7
50	0,43
80	0,27
150	0,14
280	0,1
1000	0,025

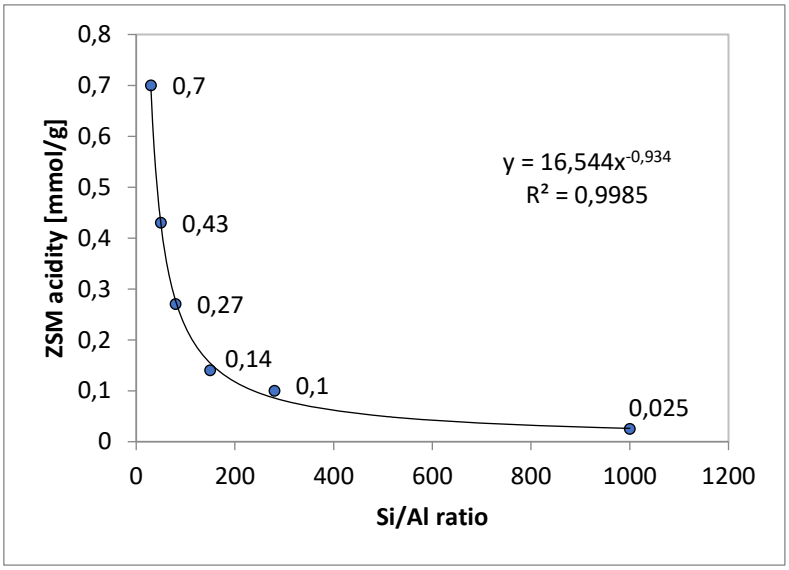


Figure 3-1. Al-ZSM Acidity vs. Si/Al ratio plot

3.2.2. Catalyst 2: B-ZSM-5 (20)

B-ZSM is structurally similar to Al-ZSM, as they share the same *Pentasil* base, but they contain different substituents. Therefore, these catalysts also experience similar mass transfer resistances. According to Cornaro (1989), the effect of substituting Boron (B), instead of Al, is a reduction in the acidity. There were no explicit values of the acidity stated in literature; however, ZSM based catalysts are known to be *Bronsted* acidic (Trong On *et al.*, 1995). Therefore, a reliable assumption is that this catalyst is much more (>>) acidic compared to the Tungsten trioxide (WO_3/SiO_2) catalyst. Further to this, Cornaro (1989) also found that B-ZSM catalysts are more active toward DBS isomerization reactions. Therefore, Boron substituted catalysts appear to create a more representative catalyst.

3.2.3. Catalyst 3: Al-MCM-41 (100)

This catalyst is presented purely for comparative purposes, as the physical and chemical properties of Al-MCM are very similar to B-MCM catalysts. Trong On *et al.* (1995), provides the following relative acidity scale,

$$\text{MCM-41} < \text{B-MCM-41} \leq \text{Al-MCM-41} \ll \text{B-ZSM-5}$$

In addition, B-MCM-41 and Al-MCM-41 were noted to have almost equal acidic strength (Trong On *et al.*, 1995). Literature does not contain explicit acidity values for B-MCM-41; however, data is available for Al-MCM-41. La-Salvia *et al.* (2017) measured the acidity data in Table 3-3, which was used to construct Figure 3-2. The acidity for an Al/Si ratio of 100 was then estimated to be 0.2779mmol/g.

Table 3-3. Al-MCM-41 Acidity Data (La-Saliva *et al.*, 2017)

Si/Al Ratio	Al-MCM Acidity [mmol/g]
50	0,36
15	0,51
10000	0,06

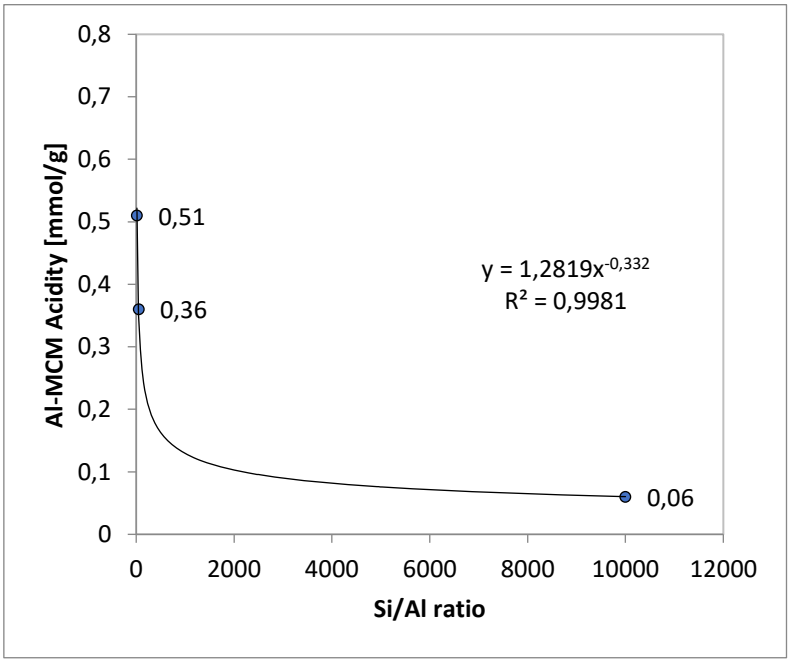


Figure 3-2. Al-MCM-41 Acidity vs. Si/Al ratio plot

3.2.4. Catalyst 4: B-MCM-41 (100)

Noting that B-MCM-41 and Al-MCM-41 have almost equal acidic strength (Trong On *et al.*, 1995), the acidity can be taken as 0.2779mmol/g for a substitution ratio of 100. Furthermore, the *pore size* (41µm) is much larger compared to the other available catalysts; therefore, the mass transfer limitations would be minimized. The acidity of B-MCM-41 is practically the same as that of the Tungsten trioxide (WO_3/SiO_2) catalyst (1.97% greater), and its pore size is sufficiently larger than the other catalysts investigated by Sundaramurthy (2000); therefore, B-MCM-41 catalyst is most representative of Tungsten trioxide (WO_3/SiO_2) catalyst, for all the catalysts investigated. The data obtained by Sundaramurthy (2000), for the B-MCM-41 catalyst is provided in Table 3-4 and will be used to develop a simple *power-law kinetic model* to describe the Isomerisation of 1-Hexene on the Tungsten trioxide (WO_3/SiO_2) catalyst.

3.3. 1-Hexene Isomerisation Model Development

3.3.1. Isomerisation Literature Data

Table 3-4 contains literature data describing the final product distribution for the Isomerisation of 1-Hexene, with B-MCM-41 (100) catalyst. This product distribution data did not specify the amount of catalyst and pure feed flowrate; however, the *Weighted hourly space velocity* (WHSV) was provided. WHSV is defined as the ratio of *hourly feed flowrate* to the *mass of catalyst in the system*. Considering that the 1-Hexene metathesis data was measured for a pure 1-Hexene flowrate of 0.26072g/min, this value was used to determine the equivalent mass of catalyst that would ensure a WHSV of $3hr^{-1}$. This results in a mass of catalyst approximately equal to 5.2145g. The WHSV ratio must be adhered to; otherwise, the model would under-predict or over-predict the identified rate constants, compared to its *true value*. This would result in inaccurately identified parameters, due to the improper description of the kinetic behaviour.

From Table 3-4 it is clear that SR products do not significantly form; therefore, this confirms the fact that the B-MCM-41 catalyst contains more Lewis acid sites and is the most representative of Tungsten trioxide (WO_3/SiO_2). This supports the assumption that *no branched alkenes are produced*; therefore, the data for 573K and 623K contains a column with the *normalized product distributions*, excluding the small percentage of SR products.

Table 3-4. 1-Hexene Isomerisation Data (Sundaramurthy, 2000)

Product	**Product Distribution (wt%) at different temperatures**						**Formation Reaction Type**
	473K	**523K**	**573K**		**623K**		
1-Hexene	82	41,3	25,4	25,8	14,8	15,2	-
2-Hexene	17	49,6	61,2	62,1	66,2	67,8	Double Bond Shift
3-Hexene	1	9,1	12	12,2	16,6	17,0	Double Bond Shift
2-methyl pentene	0	0	1,4	-	2,4	-	Skeletal Rearrangement

Experimental Conditions: WHSV = 3[1/hr], $F_{1\text{-}HEX}$ = 0.26072[g/min] and M_{CAT} = 5.2145[g]

3.3.2. Isomerisation Network and Model Equations

Tu *et al.* (2016) proposed the reaction network in Appendix B1 to describe the isomerization of 1-Hexene. The 1-Hexene isomerization network consists of *three reversible reactions*. Three reversible reactions imply that there would be *six unknown rate constants* (three forward rate constants and three reverse rate constants). Considering that there are only three model equations, it would be very difficult to find a *convergent*, but *uniform*, set of rate parameters that meet the final product distribution, at the different temperatures. In other words, there could exist numerous convergent solutions (*degree of freedom* (DOF) *underspecified*), if the values are not constrained in some way. This issue was overcome by using the *activity-based equilibrium constant*, predicted using *Aspen Plus* with the *Peng-Robinson* (PR) EOS. The activity-based equilibrium constant is independent of the type of catalyst; therefore, it can be used to re-express the *reverse reaction rate* in terms of

the *known* equilibrium constant and *unknown* forward reaction rate. Therefore, the system is now *fully specified*, and can be solved to yield a *finite solution set.* The equilibrium constants that were used are provided in Table 3-5, and fitted equilibrium constant models are provided in Figure 3-3.

The type of model, i.e., *power-law*, was selected after the consideration of competitive species adsorption on the catalyst. Other models, such as the *LHHW* type, is more useful when there is a competitive adsorption effect on a catalyst surface. However, the hexenes are *initially* present in a much higher proportion than the metathesis products; therefore, sufficient hexenes would initially be adsorbed to overcome the inhibitive effect of other species adsorption. Therefore, the isomerization of hexenes, on the B-MCM-41 catalyst, is expected to occur with almost no competitive effects, which allows for the use of a simple *power-law kinetic model*.

Table 3-5. Predicted Equilibrium Constants for Isomerisation Reactions (Aspen Plus V8.6)

Temperature	**473K**	**523K**	**573K**	**623K**
$\mathbf{K_{1e}}$	7,1455	5,5684	4,5186	3,7832
$\mathbf{K_{2e}}$	3,4336	2,6938	2,2009	1,8553
$\mathbf{K_{3e}}$	0,2309	0,2340	0,2372	0,2405

Subscripts: $\mathbf{K_{xy}}$ where **x** refers to **Reaction number** in Appendix B1, and **y** refers to **equilibrium** state.

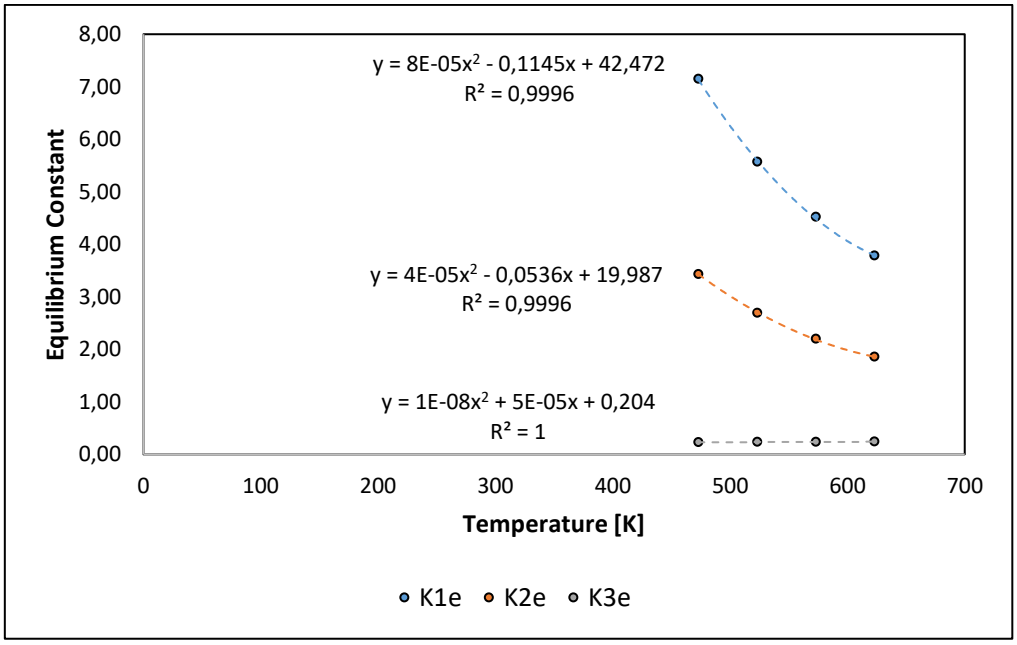


Figure 3-3. Equilibrium Constant Regression Models (Aspen Plus)

3.3.3. Rate Constant Optimization

The *rate constant optimization algorithm* is provided in Appendix B2, with the corresponding *Matlab* simulation shown in Appendix E1. The algorithm is based on the fact that the *optimal rate parameters* will result in the *lowest sum* of deviations (*e1, e2 & e3*) from the *target product distribution*, i.e., provided in Table 3-2. *e1, e2 & e3* respectively refers to the deviation of *1-Hexene*, *2-Hexene* and *3-Hexene* from its value in the *target product distribution*. The algorithm tests different combinations of *theoretical rate constants*, in different *ranges*, and finds the combination that yields predicted outlet mass fractions *closest* to those in Table 3-4, i.e., the rate constant yields the *lowest sum* of deviations *e1, e2 & e3*, for a specific temperature. The results of this *search/ optimization algorithm* are provided in Table 3-6. Table 3-7 contains the resulting deviations from the *target product distributions,* of Table 3-4, when the *optimal parameters* are used in the model equations. More accurate rate parameters could have been obtained by reducing the *granularity* of the *theoretical rate constant* search vectors; however, all deviations listed in Table 3-7 are below 1% which is well below the *experimental uncertainty* that is likely greater than 1%. Furthermore, the quality of fit is described in Figure 3-4, which shows that the *predicted data points (x)* are effectively equal to the respective *experimental data points (y)*, as they all lie on the line $y = x$ *(45˚ line)*. Therefore, the *optimal rate constants* obtained are sufficient for the purpose of modeling the 1-Hexene isomerization reaction.

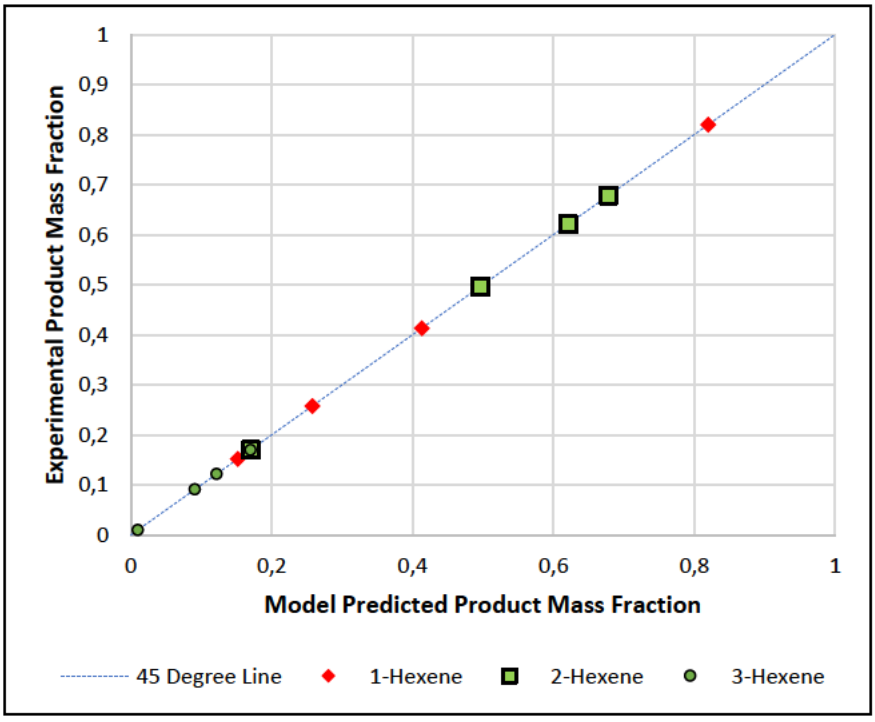


Figure 3-4. Observed vs Predicted Mass Fractions Parity Plot

The optimal reaction profiles, i.e., those that are obtained using the identified rate parameters of Table 3-6, are provided in Appendix B3. Figures B-1 to B-4, in Appendix B3, clearly show that the optimal rate parameters allow the system to achieve the *target (experimental) product distributions*, indicated in Table 3-4, while maintaining the WHSV and respective temperatures used in the literature experiments of Sundaramurthy (2000).

Table 3-6. Identified Isomerisation Rate Parameters that Minimize Error in Predicted Mass Fraction

Rate Constant (Optimal)	**Predicted Rate Constants [m^3/kg/s] at different Temperatures**			
	473K	**523K**	**573K**	**623K**
k1	7.510E-05	3.694E-04	7.653E-04	4.204E-03
k2	4.174E-06	5.327E-05	7.418E-05	1.939E-03
k3	2.469E-06	7.204E-05	1.735E-04	4.020E-03

Table 3-7. Final Mass Fraction Deviations, for the Optimal 1-Hexene Isomerisation Rate Constants

Product	**Model-predicted Absolute Mass Fraction Deviations [frac.] at different Temperatures, for Optimal Rate Constants**			
	473K	**523K**	**573K**	**623K**
1-Hexene	8,505E-06	3,598E-04	5,024E-05	3,788E-05
2-Hexene	4,695E-05	2,840E-04	2,879E-05	7,411E-06
3-Hexene	1,007E-04	8,514E-05	2,532E-04	4,215E-06

3.3.4. Kinetic model: Linear Regression & Statistical Results

Table 3-8. Linear Regression & Statistical Results for Optimal Isomerisation Parameters

Statistical Parameters	**Reaction 1**	**Reaction 2**	**Reaction 3**
$\mathbf{R^2}$	0.9712	0.9072	0.9578
$\mathbf{F_{value}}$	67.41	19.55	45.34
$\mathbf{b_o}$	6.323	10.67	15.86
$\mathbf{b_{o\ Error}}$	1.694	4.585	3.734
$\mathbf{P_{o\ value}}$	0.065	0.145	0.051
$\mathbf{b_1}$	-7 501.32	-10 935.63	-13 563.00
$\mathbf{b_{1\ Error}}$	913.62	2473.05	2014.30
$\mathbf{P_{1\ value}}$	0.015	0.048	0.021
Activation Energy [kJ/mol]	62.37	90.92	112.76
Frequency Factor [m^3/kg/s]	556.46	43 105.25	7 697 452.59

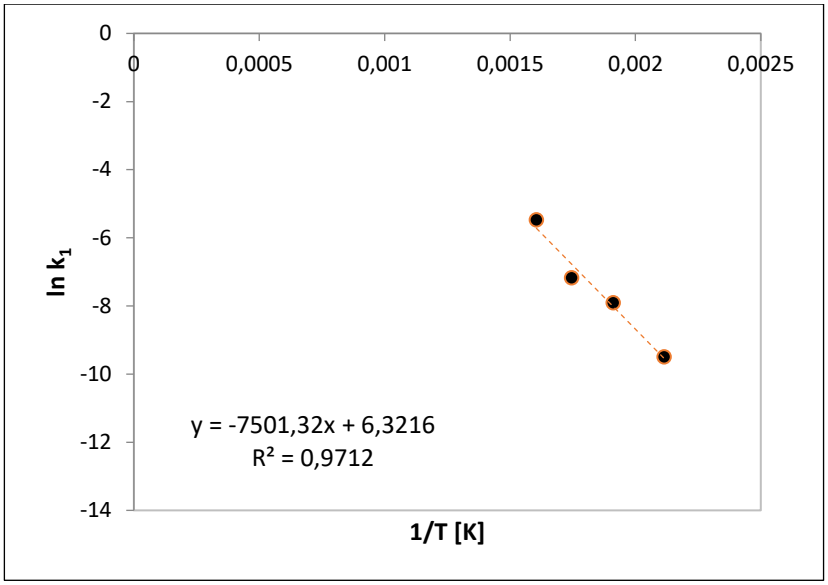


Figure 3-5. Arrhenius Plot for Reaction 1

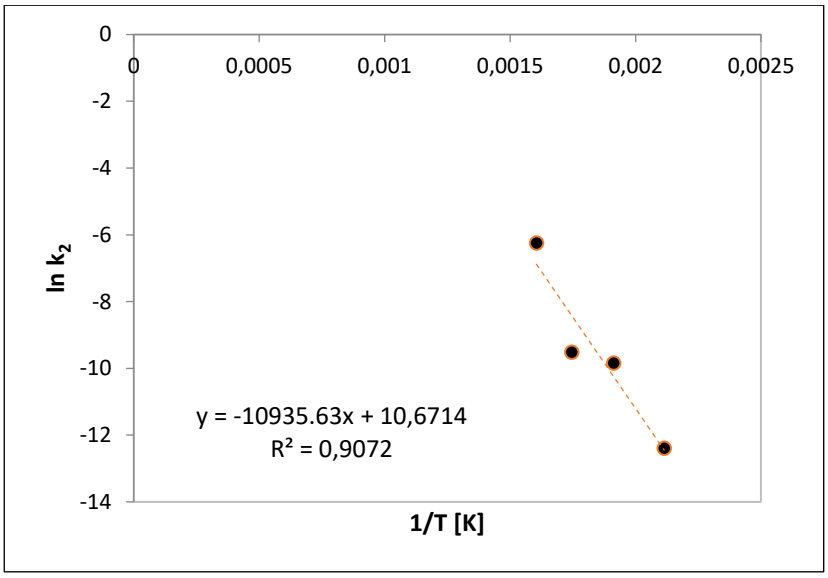


Figure 3-6. Arrhenius Plot for Reaction 2

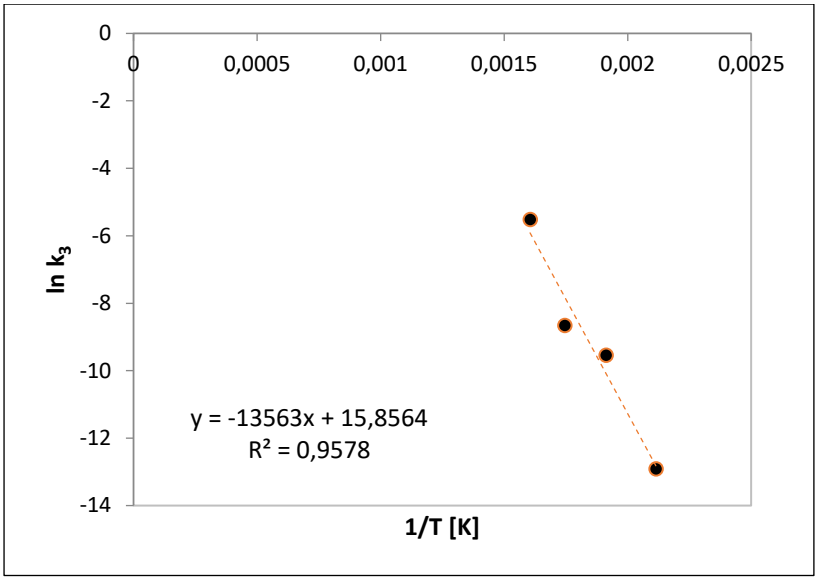


Figure 3-7. Arrhenius Plot for Reaction 3

3.4. Discussion

The identified (optimal) reaction rate constants, along with the associated temperatures, are provided in Table 3-6. The *Arrhenius plot* was constructed for each reaction's *optimal rate constant* and *temperature* combinations, after the required data transformations were made, according to Section 2.13.3. The *activation energy* and *frequency factor* for each reaction is provided in Table 3-8 and was obtained using the *Arrhenius plot slope* and *y-intercept* with the method illustrated in Section 2.13.3. The *python statsmodels* package was used to fit the regression models, for each reaction. The results of the regression procedures are provided in Table 3-8, and graphically represented in Figures 3-5, 3-6 and 3-7.

3.4.1. Model Trends

A peculiar trend is observed in Figure 3-5 to 3-7. Each data point appears to show residuals that increase in the same direction, as the respective point in the previous plot. This makes it seem like the model should be composed of numerous line segments of different slopes, i.e., a *multi-step mechanism*, instead of a single linear model. As illustrated in Section 2.14, it is worth noting that some reactions go through different mechanisms due to a change in the controlling step. This can be associated with catalyst surface structure, temperature, and other factors; therefore, these different slopes could be indicative of possible mechanism changes. However, there are too few data points between each existing point, to definitively conclude that such a *multi-step process* exists. Therefore, a single linear model was selected for each reaction's *optimal rate parameter* dataset. This linear model will be thoroughly examined for *statistical significance* to prove that even if such a *multi-step mechanism* exists, its effect would be statistically negligible. The analysis proving *statistical significance* is detailed below.

3.4.2. Coefficient of Determination (R^2) Analysis

The *R^2 coefficient* values are provided in Table 3-8, and it indicates the amount of variability in the data that is captured/ explained by the model. Therefore, the regressed reaction models perform quite well, as the lowest *R^2 coefficient* was 0.9072. This implies that for the worst-case fit, i.e., *Reaction 2*, the model captures 90.72% of the variability in the data. The *R^2 coefficient* values for *Reactions 1* and *3* are comparatively higher and have values of 0.9712 and 0.9578, respectively. The *unexplained variability* is as a result of one or more features that have not been accounted for; however, the *linearity* of the relationship in the data is confirmed by the high *R^2 coefficient* values (>0.9).

3.4.3. P-value Analysis

The *P-value* indicates if a specific variable is *statistically significant*. *Statistical significance* indicates if the relationships observed in the *sample* can be generalized to the *population*, and is determined based on a hypothesis test, with the *null hypothesis* being that the regressed weights are zero, and the *alternate hypothesis* being that the weights are not zero. In this case, the *P-value* refers to the *significance* of the model *weights*, i.e., the *slope (b_1)* and *y-intercept (b_0)*. Table 3-8 indicates that the *slopes (b_1)* for all reactions are *significant*, as all *P-values* are less than 0.05; therefore, there is enough evidence to *reject* the *null hypothesis*, i.e., the regressed *slopes (b_1)* are not equal to zero. Therefore, the *independent variable (1/T)* is associated with changes in the *dependent variable (ln k)*, at the population level.

The *P-values* for the *y-intercepts (b_0)* do not matter in the same way as those for the *slopes (b_1)*. In terms of theory, the *y-intercepts (b_0)* represent the rate constant at infinite temperature, which is an impossible notion. Furthermore, the *y-intercepts (b_0)* lie outside the range of the data sample; therefore, the linear relationship may not necessarily hold, or it may change. Therefore, statistical significance does not hold as much importance for the *y-intercepts (b_0)*; however, the *y-intercepts (b_0)* must still be included (whether significant or not) as these intercepts force the regression residuals to have a *zero mean*, i.e., no residual bias.

3.4.4. F-value Analysis

The *F-value* combines the effect of all variables together and provides an overall measure of a model's *statistical significance*; therefore, if *any weights* in a model are *significant*, then the *F-value* would be one that indicates *significance*. The *F-values* presented in Table 3-8, are from the *F-test* of overall *significance*. This *hypothesis test* indicates if the *regressed model* is better than one which contains *no independent variables*, i.e., only an intercept. Therefore, the *null hypothesis* is that the model with *no independent variables* fits the data *as well as* the *regressed model*, and the *alternate hypothesis* is that the regressed model *better* fits the data. Given that the *P-values* of the *slopes (b_1)* are all *significant*, this allows for the conclusion that the *regressed models* better fit the respective datasets, when compared to the *intercept-only model*. This further confirms the observed correlations indicated by the *coefficients of determination (R^2)*.

4. Group Contribution Modeling Methodology

4.1. Background

Froment (2005) proposed a *single event* (*SE*) kinetic modeling technique and implemented it in complex hydrocarbon systems. The *SE* technique possessed the advantage of being independent of typical feedstock compositions and reactor configurations, whilst maintaining sufficient predictive accuracy. Such advantages (independencies) occur as the *elementary step* rate coefficients are modeled based on *transition state theory* and *statistical thermodynamics* (Froment, 2005). Kinetic model accuracy, coupled with independence of the feedstock composition, is important for the design and scale-up of typical hydrocarbon processing plants.

Refinery streams can have an extremely variable composition; therefore, conventional kinetic models result in an excessive number of reactions, thus rate coefficients. The main advantage of the *SE* approach involves a reduced computational intensity, which is associated with the fact that the number of elementary steps possible is always less than the number of molecules present in a single reaction (Froment, 2005).

The issue with directly applying the *SE* approach to the investigated metathesis reaction, lies with the mechanism of the reaction. From Section 2.1 and 2.5, the *overall reactions* pass through *two activated complexes/ transition states,* i.e., a *multi-step mechanism*, before product formation. Therefore, the overall system falls under *limitation case (1)* of the Transition state theory (*TST*), as explained in Section 2.11. This implies that the development of the overall rate coefficient cannot be carried out in terms of the *TST*; as such, the *SE* approach cannot be readily applied. Therefore, a modification is proposed here, i.e., a modeling approach that draws on the *concept* of the reaction rate constant being directly related to the *alkyl groups* involved in the formation of the transition state complex.

4.2. Network Generation

The reaction network was generated using a *Matlab* simulation. This was achieved using a *vectorized notation* to represent the hydrocarbons in *Matlab*. The following assumptions were applied to develop the metathesis reaction network,

1. The maximum allowable carbon chain length of metathesis products is '*10*'. This is supported by the experimental data in Appendix A as very small amounts of hydrocarbons with length above '*10*' formed. Furthermore, all metathesis reactants and products formed are assumed to be of the *linear alkene variant*. *Branching* generally occurs with longer chain hydrocarbons, and is not very significant due to the amounts of product that form in the experimental data set (Appendix A).
2. Isomerisation is more common for short-chain alkenes ($<C_7$), compared to long-chain alkenes; therefore, the components most likely to isomerize would be the *butenes*, *hexenes* and *pentenes*. However, only *1-hexene* is assumed to isomerize to form *2-hexene* and *3-hexene*. This is a logical assumption as the feed only consists of *1-hexene;* as such, the concentration-based driving force for *1-*

hexene isomerization will be most significant. The literature data in Appendix A supports this, as very small amounts of hydrocarbons above C_{11} formed, for the higher temperature datasets.

3. The reaction temperature (~500°C) results in a gas phase reaction. Therefore, all components that have a normal boiling temperature (T_b) below the reactor temperature (T_{rxr}), i.e., all chemical species, will participate in the gas-phase metathesis reactions.

These assumptions are *sufficient* to ensure that the generated reaction network emulates the literature experimental results, from Appendix A. An important step in ensuring generated network validity is to consider both the *type (chain-length)* and *amount (moles)* of all products that form; such that, if a negligible amount of a specific product forms, then such products can be removed from the generated/ predicted network. This gives rise to the *necessary assumption*, i.e., there are negligible amounts of hydrocarbons with chain-length above '10' ($>C_{10}$) that forms. These assumptions will ensure accuracy in the generated network, which is important as the *identified rate parameters* must be as accurate as possible to ensure the proper conclusions about the *kinetic model* and *group contribution approach* are reached. The resulting reaction network is provided in Table C-1 of Appendix C2.

4.2.1. Network Generation: Hydrocarbon Notation and Generation Algorithm

A block diagram illustrating the general logic used to develop the *network generation script* is provided in Appendix C1. To develop a network generator, a *linear hydrocarbon* notation was developed. A typical hydrocarbon is represented, using the developed *linear hydrocarbon notation*, in Figure 4-1.

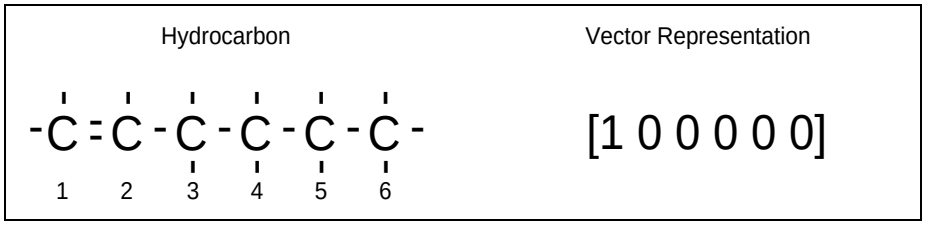


Figure 4-1. Typical Hydrocarbon Representation

As illustrated in Figure 4-1, the position where the double bond begins is denoted with a '1' in the vector representation and the remaining carbons are denoted with '0'. It is important to note that the length of the hydrocarbon chain is equal to the length of the vector representation. Based on the assumption of only linear alkenes being formed and participating in the metathesis reactions, each hydrocarbon can be uniquely identified by means of their *length* and *double bond position*. This was the basis for the development of unique identifiers in the network generation simulation, which were used to track the type of products that formed and ensured that the same reactant combinations were not considered again. It is important to note that by ensuring the uniqueness of reactant pairs, the products that form are immediately considered valid even though the same products could result from a different reactant combination.

All *validity constraints* implemented are listed below,

1. *Reactant uniqueness* – Ensured that the reactants used were not repeated in any way. Proper implementation required the development of a function that ensured the double bond numbering did not exceed half the hydrocarbon chain length. In terms of the IUPAC conventions, this translates to ensuring that the lowest possible number is allocated to the double bond position.
2. *Product length* – Ensured that the length of *product pairs* did not exceed the maximum allowable *hydrocarbon chain-length*, i.e., '10'. This is important in cases where internal olefins (alkenes) are involved, as '2' possible product pairs can form.
3. $Pi \neq Ri$ – For a *corresponding* product pair and reactant pair, this ensured that the products that formed were different from the reactants, as such reactions are invalid.

The placements of these validity constraints, regarding program flow, are clearly illustrated in the *Network generation block diagram* (Appendix C1). The algorithm was constructed such that any new products that form are also considered possible reactants. This illustrated the importance of accounting for 1-Hexene isomerization (Chapter 3), as the current reaction network, consisting of 55 Reactions, differs vastly when compared to the case where no isomerization occurs (only 1 Reaction).

4.3. Equilibrium Constant Analysis

The equilibrium constants for each reaction are provided in Table C-1 of Appendix C2, for the experimental temperatures in Table A-1. All the equilibrium constant values in Table C-1 were within the range of 0.01 to 100 (Smith, Van Ness and Abbott, 2001); this implies that *none* of the reactions show dominance in a particular direction; as such, all reactions must be considered *reversible*.

4.3.1. Equilibrium Constant Estimation Technique

The metathesis reaction equilibrium constants were estimated using the method outlined in Appendix C3, which is based on the work of Smith, Van Ness and Abbott (2001). Smith, Van Ness and Abbott (2001) did not provide all the required properties for some isomers and components. Therefore, the missing properties were either obtained from the *NIST*® database, or manually estimated using the *Joback-Reid group contribution method* which is illustrated in Appendix C3. The sources from which the unknown properties were determined are also indicated in Table C-2 of Appendix C3.

In cases where the *Joback-Reid method* was used to determine the species *heat capacities*, the equations obtained were refitted to the algebraic form required by Smith, Van Ness and Abbott *(SVA)* (2001). This resulted in the following transformation,

$$\underbrace{\frac{C_p^o}{R} = W + XT + YT^2 + ZT^3}_{\boldsymbol{Joback - Reid\ Form}} \quad \rightarrow \quad \underbrace{\frac{C_p^o}{R} = A + BT + CT^2 + \frac{D}{T^2}}_{\boldsymbol{SVA\ Form}}$$

The method used to perform this transformation is described in Appendix C3. The properties illustrated in Table C-2 were then implemented as an extension of the network generation algorithm of Appendix C1, and the equilibrium constants were calculated at the desired temperatures.

4.4. The Metathesis Mechanism: Path Competitiveness

Consider Figure C-1 and C-2, both represent the same reaction, i.e., Reaction 3 from Table C-1, but occur with different *Alkylidenes*. This indicates that different combinations of *Alkylidenes*, could result in the same *overall reaction* via two possible paths. Therefore, there is a competitive nature that must exist between these two paths, as all *Alkylidenes* are assumed to be available.

However, it is assumed that the *least bulky path*, i.e., the one where fewer carbon atoms are involved in the *Alkylidene*s, is the controlling path. This assumption is based on the fact that smaller *Alkylidenes* are likely to form and react faster than larger variants. Furthermore, Section 2.1.3 illustrates that smaller *Alkylidenes* would provide reduced *steric hinderance* during the attachment of constituents when forming the *metallacyclobutane* complex. Therefore, applying this assumption to Figure C-1 and C-2 implies that the path shown in Figure C-2 would be considered the *dominant path*. The selection of the *dominant path* for this example, is clarified in Table 4-1.

Table 4-1. Possible Paths for Reaction 3

	R_1	P_1	M_1	R_2	P_2	M_2	Total M_1 & M_2 carbons
Path 1	CH_2=CH-C_4H_9 [1-Hexene]	CH_2=CH-CH_3 [1-Propene]	M=CH-CH_3	CH_3-CH=CH-C_3H_7 [2-Hexene]	C_3H_7-CH=CH-C_4H_9 [4-Nonene]	M=CH-C_4H_9	7
Path 2	CH_2=CH-C_4H_9 [1-Hexene]	C_3H_7-CH=CH-C_4H_9 [4-Nonene]	M=CH-C_3H_7	CH_3-CH=CH-C_3H_7 [2-Hexene]	CH_2=CH-CH_3 [1-Propene]	M=CH_2	5

Where, M_1 and M_2 refer to the *Alkylidenes.*

From Table 4-1, it is clear that *Path 2* yields the *least bulky Alkylidenes*, as the total M_1 & M_2 carbons are the lowest. Also note that the R_1-P_1 and R_2-P_2 pairs differ depending on the *Alkylidenes* involved in the reaction. The *Alkylidenes* (M_1 & M_2) in Table C-6 were all selected from the *dominant path*, which was determined as described in this section.

In some cases, the total M_1 & M_2 carbons were the same in both paths. Consider the example in Table 4-2, the total M_1 & M_2 carbons are the same in both paths. In such cases, the path with the most commonly occurring *Alkylidenes* are selected. The most commonly occurring *Alkylidenes*, i.e., M=CH_2 or M=C_4H_9, are those that form

during the *1-hexene self-metathesis reaction*, as the experimental feed consists of pure 1-hexene. This creates a large driving force for the self-metathesis of 1-hexene, which will ensure that a large proportion of $M{=}CH_2$ and $M{=}C_4H_9$ Alkylidenes are available. Therefore, *Path 1* from Table 4-2 would be selected as the *dominant path*.

Table 4-2. Possible Paths for Reaction 22

	R_1	P_1	M_1	R_2	P_2	M_2	Total M_1 & M_2 carbons
Path 1	CH_3-CH=CH-C_3H_7 [2-Hexene]	CH_2=CH-CH_3 [1-Propene]	M=CH_2	CH_2=CH-C_2H_5 [1-Butene]	C_2H_5-CH=CH-C_3H_7 [3-Heptene]	M=CH-C_3H_7	5
Path 2	CH_3-CH=CH-C_3H_7 [2-Hexene]	C_2H_5-CH=CH-C_3H_7 [3-Heptene]	M=CH-C_2H_5	CH_2=CH-C_2H_5 [1-Butene]	CH_2=CH-CH_3 [1-Propene]	M=CH-CH_3	5

Where, M_1 and M_2 refer to the *Alkylidenes.*

4.5. Group Contribution Methodology

In Appendix C4 selected reactions were illustrated in terms of the *Chauvin mechanism* of Section 2.1 as it more clearly represents the *metallacyclobutane* complex. This was done for all the reactions and led to the observation that only *five unique constituents* exist in the transition state complexes/ alkylidenes, namely,

1. n-pentyl: >CH-C_4H_9
2. n-butyl: >CH-C_3H_7
3. n-propyl: >CH-C_2H_5
4. n-ethyl: >CH-CH_3
5. n-methyl: >CH_2

In this work, each constituent (group) is hypothesized to have a *unique effect on the formation rate* which influences the rate at which the reaction will occur. Therefore, the formation of the transition states is considered the *controlling step* in the metathesis reaction. The size of the *transition state constituents* is a key factor when it comes to determining the overall rate of the reaction, as *larger constituents* will take a longer time to form a transition state compared to *smaller constituents* – This is based on Section 2.1.3. This creates the framework for determining whether a *constituent-group* will affect the modeling of the rate coefficient. *Hypotheses* for the definition of the rate constant according to the substituent groups are presented, discussed, and theoretically evaluated in the following sub-sections.

4.5.1. Hypothesis 1

This *hypothesis* postulates that, the rate of a certain metathesis reaction should be controlled by the *rate of attachment* of the *largest constituent* present in the single *bulkiest transition state* (*metallacyclobutane*) complex, within the *dominant path*. If such a methodology is adopted then multiple reactions would share the same rate coefficient, because there is a high likelihood that the reaction-controlling (bulkiest) *transition state complexes* would share the same largest constituent. This is attributed to the fact that there are only *five unique constituents*.

The problem with this hypothesis presents itself when the rate coefficient of any reverse reaction is considered. Consider the rate expressions for the *overall metathesis reaction*, in Appendix C5-1; in keeping with *hypothesis 1*, it would imply that the *reverse reaction* is also subject to the same rate constant definition, as that of the *forward reaction*. Therefore, the resulting reaction equilibrium constant would have a value of *one*. Table C-1 illustrates that this holds for *one reaction*, i.e., Reaction 37, but all other reactions do not show equilibrium constants that are approximately equal to *one*. Therefore, *hypothesis 1* requires modification for other reactions in which both directions, i.e., forward and reverse, are significant but no single path dominates.

The rejection of *hypothesis 1* currently occurs at the *overall* reaction level; therefore, it must be examined at the *elementary* level of the reaction, before it can be completely rejected. Consider the rate expressions for the *elementary metathesis reaction steps* in Appendix C5-2; the main rate coefficients that need to be defined are k_1, k_4, k_5 and k_8. These rate coefficients all account for the reaction of an *Alkylidene* (M_1 & M_2) with an *Overall chemical component* (P_1, R_1, P_2 & R_2).

Firstly, *hypothesis 1* is modified to account for the fact that elementary level is now considered. The *modified hypothesis 1* requires that the rate constants be defined according to the largest constituent, in the respective *transition state* of each *elementary step*. Therefore, k_1 would have the same definition as k_4, and k_5 would have the same definition as k_8. The reason for this is that each of these rate constant pairs operate around the same respective transition state (TS) complexes, i.e., k_1 & k_4 define rates to form TS_1 and k_5 & k_8 define rates to form TS_2. So, the same largest constituent will be present, as the same transition states are involved.

To determine if this is sensible, the way in which the transition state (TS) complex forms must be understood. According to Section 2.1, the TS complex formed by the attachment of two *external (incoming)* constituents to the existing *Alkylidene*. The *Alkylidene* already contains a certain *internal constituent group*, and if this were to be the *largest constituent* then the rate constant would be incorrectly defined as a constituent that is already attached. Therefore, this proves that *hypothesis 1* cannot be completely true, as the *largest constituent* cannot always be selected to define the rate constant.

4.5.2. Hypothesis 2

Hypothesis 2 was developed to correct the issues encountered with the *modified hypothesis 1*. The rate constant is now defined as the *largest constituent* that is not present in the reacting *Alkylidene*, i.e., the *largest constituent* in the *external species* (R_1, R_2, P_1 or P_2). Therefore, at the *elementary level*, only the *largest external constituent* group that attaches to the *Alkylidene* is used to define the rate constant (k_1, k_4, k_5 and k_8). This is a more theoretically sound hypothesis as it is based on the attachment of groups that are not already present in the *Alkylidene*. Therefore, k_1 would not necessarily have the same definition as k_4, and k_5 would not necessarily have the same definition as k_8. The reason for this is that even though each of these rate constant pairs operate around the same respective transition state (TS) complexes, i.e. k_1 & k_4 define rates to form TS_1 and k_5 & k_8 define rates to form TS_2, they define the formation of the same transition state from different *Alkylidenes* and *Species*, i.e., $k_1 = f(M_1, R_1)$ & $k_4 = f(M_2, P_1)$ and $k_5 = f(M_2, R_2)$ & $k_8 = f(M_1, P_2)$.

Another observation from the QSSA model rate expressions, in Appendix C5-2, is that the group contribution rate constants (k_1, k_4, k_5 and k_8) are always multiplied by either the concentration of Alkylidene M_1 or M_2. Considering that this is effectively the concentration of an active site on the catalyst surface, it cannot be determined through calculation. Therefore, these active site concentrations must be lumped with the group contribution rate constants as shown in Appendix C5-2. The lumped relations are shown below,

$$k_{1,M_1} = k_1 C_{M_1}\ ,\ k_{5,M_2} = k_5 C_{M_2}\ ,\ k_{4,M_2} = k_4 C_{M_2}\ \ \&\ \ k_{8,M_1} = k_8 C_{M_1}$$

Where, C_{M_i} refers to Alkylidene or active site concentration, and k_i refers to the group contribution rate constant. The *group definitions* of rate constants k_{1,M_1}, k_{4,M_2}, k_{5,M_2} and k_{8,M_1}, according to *hypothesis 2*, are provided in Table C-6.

4.5.2.1. Equilibrium Constant Effect

Hypothesis 1 clarified the fact that the reaction must be examined at the *elementary level*, to reach a conclusion about the group contribution methodology. However, this only addressed the definition of some *Quasi-steady state approximated* (QSSA) rate constants, i.e., k_{1,M_1}, k_{4,M_2}, k_{5,M_2} and k_{8,M_1}. The remaining rate constants, i.e., k_2, k_3, k_6 and k_7, must also be defined in some manner. The complexity arises as these rate constants deal with the *dissociation* of the Transition state complexes (TS_1 & TS_2).

This *transition state dissociation* (*breaking apart*) process is energy-driven and estimating these rates involves determining the activation energy for the formation of the transition state complex, from the respective *Alkylidene* and *reacting Species* pairs. The rate constants k_2 & k_3 and k_6 & k_7 need to be quantified in order to determine the validity of the group contribution methodology. Each of these rate constants deal with the dissociation of the Transition state to form the *products* (k_3 & k_7), or *reform* the *reactants* (k_2 & k_6). Since these rate constants deal with the dissociation of the *Transition state*, it should be predicted using a method closely linked to *Transition state theory*. The only way to independently estimate the *unknown rate constants* (k_2, k_3,

k_6 and k_7) is using the *Gibbs formation energies*, as shown in Section 2.11.2. However, given that the required species formation data is not available for the *Transition state complexes* and *Alkylidenes*, the required *Gibbs energies* would have to be determined using quantum chemistry simulations or statistical thermodynamics.

This issue with using quantum chemistry simulations to estimate these *energetic parameters* is that its *correctness* will have a direct impact on the accuracy of the kinetic modeling. This occurs as the optimal values of the *defined rate constants* (k_{1,M_1}, k_{4,M_2}, k_{5,M_2} and k_{8,M_1}) are dependent on *Activation energies* of the *elementary steps* of a particular reaction. Reason being, the values of k_1, k_2, k_3, k_4, k_5, k_6, k_7 & k_8 are linked by means of the respective *elementary step metathesis equilibrium constants* as shown in Figure 4-2 and Appendix C5-2. Therefore, the value used for k_2, k_3, k_6 and k_7 will directly affect the regression procedure and results for the unlumped group contribution rate constants (k_1, k_4, k_5 and k_8), which will affect the conclusion reached in this work. The following subsection proposes and assesses a way to overcome this issue.

Individual Metathesis Reaction Equilibrium Relations

$$K_A = \frac{k_1}{k_2} \quad K_B = \frac{k_3}{k_4} \quad K_C = \frac{k_5}{k_6} \quad K_D = \frac{k_7}{k_8}$$

Elementary Step Metathesis Equilibrium Relations

$$K_{e,1} = (K_A \times K_B) \quad K_{e,2} = (K_C \times K_D)$$

Multistep Overall Metathesis Equilibrium Relation

$$K_e = K_{e,1} \times K_{e,2} = \left(\frac{k_1}{k_2} \times \frac{k_3}{k_4}\right) \times \left(\frac{k_5}{k_6} \times \frac{k_7}{k_8}\right)$$

Figure 4-2. Equilibrium Relations at different Levels of the Reaction Mechanism

4.5.2.2. Proposal

This section proposes that the *overall equilibrium constant* is related to the *transition state dissociation process*, by means of an analogous forward-reverse dissociation process. Therefore, the dissociation of the Transition state to form the respective product is analogous to a *forward reaction*, and the dissociation of the Transition state to form the respective reactant is analogous to a *reverse reaction*. This is represented in Figure 4-3, where k_3 & k_7 are the analogous *forward dissociation* rates and k_2 & k_6 are the analogous *reverse dissociation* rates, for TS_1 and TS_2 respectively.

More importantly, Figure 4-3 also proposes that the ratio of the *analogous forward* to *reverse dissociation* rate constants, i.e., $\frac{k_3}{k_2}$ & $\frac{k_7}{k_6}$, are approximately equal. This proposal will be evaluated in the next section. Considering that the breakage of bonds controls the rate of transition state dissociation, and that specific bonds break depending on the direction of the dissociation reaction, the *analogous forward* (k_3 & k_7) and *reverse* (k_2 & k_6) *dissociation* rate constants should be directly related to the strongest bonds that break in their respective transition states. The bond dissociation energy relationships that possibly exist can be used to elucidate the extent to which this proposal is valid.

Elementary Step 1: Reactions around TS_1

$$M_1 + R_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TS_1 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} M_2 + P_1$$

Elementary Step 2: Reactions around TS_2

$$M_2 + R_2 \underset{k_6}{\overset{k_5}{\rightleftharpoons}} TS_2 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} M_1 + P_2$$

Proposed Relation

$$\frac{k_{forward}}{k_{reverse}} = \frac{k_3}{k_2} \approx \frac{k_7}{k_6}$$

Figure 4-3. Proposal 1 Represented in Terms of Dissociation Rate Constants

4.5.2.3. Bond Dissociation Energies & Transition State Dissociation

Figure C-3 in Appendix C4 illustrates the *dominant* path for Reaction 1 of Table C-1, as this is a case of *self-metathesis* and there will exist only one pair of possible *alkylidenes* resulting in a single dominant path. In Figure C-3, two additional dashed lines are illustrated on the transition states (TS_1 & TS_2). These dashed lines illustrate which bonds break during the dissociation of the transition state complex, in accordance with the respective direction of dissociation. The dashed bond breakage lines are colour coded to indicate which bonds break during the *forward* (blue dashed line) and *reverse* (red dashed line) processes.

From Figure C-3 there is always one M-C bond and one C-C bond that breaks during the dissociation of a single transition state. Therefore, the rate of dissociation of the transition state must be directly related to the breakage of these bonds. From Figure 2-3, Carbon (C) bonds to the Tungsten (W) in species M; therefore, the bond strength of W-C and C-C bonds must be analysed to determine its effect on the dissociation process. The pure bond dissociation energies of interest are presented in Table 4-3.

Table 4-3. Pure Bond Dissociation Energies

Bond	**Bond Dissociation Energy**		**Reference**
	[kJ/mol]	**[eV]**	
W-C	510	5.289	Sevy, Huffaker, and Morse (2017)
C-C	607	6.291	Speight, Lange, and Dean (2005)

Note: W refers to Tungsten, and C refers to carbon

The higher the bond dissociation energy, the more energy is required to break the bond. Therefore, the bond dissociation energies in Table 4-3 appear to indicate that the Carbon-carbon (C-C) bond breakage controls the transition state dissociation process. To conclude that this is true, the effect of external groups on the C-C and W-C bonds must be determined.

Table 4-4. Carbon-Carbon (C-C) Bond Dissociation Energies in Different Chemical Species

Bond	**C-C Bond Dissociation Energy**		**Reference**
	[kJ/mol]	**[eV]**	
$H_3C\text{-}CH_3$	376	3.897	Cao and Yuan (2003)
$(H_3C)_2HC\text{-}CH_3$	335	3.472	Speight, Lange, and Dean (2005)
$(H_3C)_2HC\text{-}CH(CH_3)_2$	282.4	2.927	Speight, Lange, and Dean (2005)

The data presented in Table 4-4 was obtained using the *enthalpy of formation* (ΔH_f) of the organic radicals used to form the respective chemical species. It does appear that the C-C bond dissociation varies significantly across species; however, this method does not account for changes in conformation (spatial arrangement) or include Zero-point Energy (ZPE) and thermal correction factors. Therefore, these results are not based on the actual strength of the C-C bonds in the species; rather, it provides a measure of the expected bond dissociation energy based on thermodynamic energy conservation principles. Hunter and East (2002) investigated the true bond dissociation energy of the C-C bond in various alkanes, i.e., alkane carbon chain lengths of 2 up to 20. In all alkanes, they found that the C-C bond dissociation energy was relatively constant, at about 88 kcal/mol (368 kJ/mol) at 298K, except for the terminal C-C bonds, which were around 89 kcal/mol (372 kJ/mol), and the penultimate C-C bonds, which were around 87 kcal/mol (364 kJ/mol), respectively (Hunter and East, 2002). Hunter and East (2002) attributed this result to the stability of ethyl radicals over other primary radicals.

Alkorta and Elguero (2006) conducted a similar investigation and obtained results that validated the conclusions obtained by Hunter and East (2002). Therefore, the strength of C-C bonds that dissociate can be considered approximately equal for a specific transition state. However, there is still no certainty about whether C-C bond breakage controls the dissociation of the entire transition state, as the W-C bond strength must also be assessed.

Table 4-5. Tungsten-Carbon (W-C) Bond Dissociation Energies in Different Chemical Species

Bond	**W-C Character**	**W-C Bond Dissociation Energy**		**Reference**
		[kJ/mol]	**[eV]**	
W-$(CH_3)_6$	W-alkyl	160	1.658	Simoes and Beauchamp (1990)
W-$(Me)_2(Cp)_2$	W-alkyl	221	2.291	Simoes and Beauchamp (1990)
W-$(CO)_6$	W-carbonyl	192	1.990	Simoes and Beauchamp (1990)
W-CH_3	W-alkyl	155	1.606	Skinner and Connor (1985)
W-C_6H_6	W-arene	298	3.089	Skinner and Connor (1985)
W-CO	W-carbonyl	176	1.824	Skinner and Connor (1985)

Note: Cp refers to Cyclopentadienyls – C_5H_5

Tungsten is not present as a pure element in the W-C bond, as it forms part of a larger complex on the catalyst surface. This complex is illustrated in Figure 2-3, and possesses two W=O groups. Theoretically, the presence of oxygen groups is expected to shift electron density away from the W-C bond, which would reduce the shared electron density between Tungsten and Carbon, resulting in a weakened bond. According to Skinner and Connor (1985), the character of the W-C bond creates a distinct bond dissociation energy scale, which is described as follows,

$$BDE(W-Cp) > BDE(W-arene) > BDE(W-alkyl) > BDE(W-carbonyl)$$

The different W-C bond characters, with its associated bond dissociation energy, is provided in Table 4-5. By comparing this scale with the data in Table 4-5, it is clear that this scale is correct for the Arene and Alkyl characters, but not for the carbonyl character. However, all that this work requires is an upper bound on the W-C bond dissociation energy, which is provided by the Arene bond dissociation energy. Given that the C-C bond energy, obtained from Hunter and East (2002), is approximately constant at 88 kcal/mol (368 kJ/mol), and the W-C bond energy is upper-bounded at approximately 71.2 kcal/mol (298 kJ/mol); it becomes clear that the linear C-C bond is likely always stronger than the linear W-C bond. This observation is expected to become

more pronounced due to the steric interaction between the W=O groups and substituent interactions in the metallacycle (transition state) ring, in accordance with the findings of Section 2.1.3.

Therefore, the C-C bond dissociation must be the controlling process in the dissociation of the transition states in a single elementary reaction mechanism. Furthermore, the breaking C-C bond dissociation energies were proven to be independent of where they occur (Hunter and East, 2002); therefore, it can be treated as approximately equal for all C-C bonds in a specific reaction's transition state pair. Therefore, the developments shown here, combined with the transition state theory of Section 2.11.2, can be used to elucidate the following generalized relationship,

$$\frac{(\mathrm{C}-\mathrm{C}\text{ dissociation for step yielding }\mathrm{M}_2\text{ and }\mathrm{P}_1)_{TS1}}{(\mathrm{C}-\mathrm{C}\text{ dissociation for step yielding }\mathrm{M}_1\text{ and }\mathrm{R}_1)_{\mathrm{TS1}}} \approx \frac{(\mathrm{C}-\mathrm{C}\text{ dissociation for step yielding }\mathrm{M}_1\text{ and }\mathrm{P}_2)_{\mathrm{TS2}}}{(\mathrm{C}-\mathrm{C}\text{ dissociation for step yielding }\mathrm{M}_2\text{ and }\mathrm{R}_2)_{\mathrm{TS2}}}$$

Where, P_1 and P_2 refer to the different products, R_1 and R_2 refer to the different reactants, and M_1 and M_2 refer to different alkylidenes, as shown in Figure 4-3 and Figure C-3. This relationship exists as the C-C bond breakage was proven to be the controlling process for transition state dissociation as it has a bond dissociation energy larger than that of W-C, and it has a bond dissociation energy independent of where it appears (Alkorta and Elguero, 2006). Since the broken C-C bond dissociation energies are always similar, the ratio of the dissociation rates during the analogous forward and reverse transition state dissociation processes should be approximately equivalent, for a specific pair of transition states. Therefore, since the rates of transition state dissociation are directly related to the dissociation of the strongest (controlling) bonds, the following relationship holds,

$$\frac{k_3}{k_2} \approx \frac{k_7}{k_6}$$

It is important to note that this relationship only holds for the specific pair of transition states, that occur in a single reaction. This is a necessary step to prevent over generalizing the dissociation rate parameters. The relationship cannot be extended across all reactions as different reactants and alkylidenes (M_1 & M_2) combine to form those sets of transition states (TS_1 & TS_2); this will create a unique dissociation effect due to the substituent specific interactions with the W=O groups. Therefore, the relationship can only be individually applied to each reaction in Table C-6. Although this is contrary to expectations, given that the C-C bond was shown to have the same bond dissociation energy independently of where it occurred, the experiment validating this claim was not carried out explicitly on the *metallacyclobutane* complex – it used linear hydrocarbons (alkanes). Therefore, it is most likely that the C-C bond dissociation energy is most similar in transition states/ metallacycles that have similar substituents – like those *metallacyclobutane* C-C bonds that occur in a specific reaction's metathesis mechanism – but will be different when the transition states substituents differ – like those metallacycle C-C bonds that occur in different reaction's transition state complexes. The proposed method will ensure that the effect of transition state dissociation is accounted for in a reaction specific manner.

Incorporating this relationship into the multistep equilibrium relation, from Section 2.11.2 and Figure 4-2, provides a way to evaluate the remaining unknown factors (k_2, k_3, k_6 and k_7) in the expression. The use of the proposed approximation is illustrated in Figure 4-4, in which the unknown lumped factors involving k_2, k_3, k_6 and k_7 are re-expressed in terms of the known *overall reaction equilibrium constant* (K_e), and the *regressed group contribution rate constants* (k_1, k_4, k_5 and k_8) to be estimated. The *overall equilibrium constants* are provided in Table C-1 of Appendix C2.

Multistep Overall Equilibrium Relation

$$K_e = K_{e,1} \times K_{e,2} = \left(\frac{k_1}{k_2} \times \frac{k_3}{k_4}\right) \times \left(\frac{k_5}{k_6} \times \frac{k_7}{k_8}\right)$$

$$K_e = \left(\frac{k_1}{k_4} \times \frac{k_5}{k_8}\right) \times \left(\frac{k_3}{k_2} \times \frac{k_7}{k_6}\right)$$

Since $\frac{k_3}{k_2} \approx \frac{k_7}{k_6}$,

$$\left(\frac{k_3}{k_2}\right)^2 = \left(\frac{k_7}{k_6}\right)^2 = \frac{K_e}{\left(\frac{k_1}{k_4} \times \frac{k_5}{k_8}\right)}$$

$$\frac{k_3}{k_2} = \frac{k_7}{k_6} = \sqrt[2]{\frac{K_e}{\left(\frac{k_1}{k_4} \times \frac{k_5}{k_8}\right)}}$$

Let $\sqrt[2]{\frac{K_e}{\left(\frac{k_1}{k_4} \times \frac{k_5}{k_8}\right)}} = K$,

$$\Longrightarrow \frac{k_3}{k_2} = \frac{k_7}{k_6} = K$$

The following algebraic operations can be performed to re-express the unknown algebraic factors appearing in the QSSA rate expressions of Appendix C5-2, in terms of the known and estimated parameters,

$\frac{k_2}{(k_2+k_3)}$ and $\frac{k_6}{(k_6+k_7)}$ can be expressed in terms of K as follows,

$$(K+1) = \frac{k_3}{k_2} + \frac{k_2}{k_2}$$

$$(K+1) = \frac{k_3 + k_2}{k_2}$$

$$\Rightarrow \frac{1}{(K+1)} = \frac{k_2}{(k_2+k_3)} = \frac{k_6}{(k_6+k_7)}$$

$\frac{k_3}{(k_2+k_3)}$ and $\frac{k_7}{(k_6+k_7)}$ can be expressed in terms of K as follows,

$$\left(\frac{1}{K} + 1\right) = \frac{k_2}{k_3} + \frac{k_3}{k_3}$$

$$\left(\frac{K+1}{K}\right) = \frac{k_2 + k_3}{k_3}$$

$$\Rightarrow \frac{K}{(K+1)} = \frac{k_3}{(k_2+k_3)} = \frac{k_7}{(k_6+k_7)}$$

Figure 4-4. Incorporating Transition State Dissociation Rate Constants into the QSSA Mechanism

4.5.3. Hypothesis 3

Hypothesis 3 was developed to assess the applicability of simpler models, by avoiding the need to consider the dissociation of the transition states in the kinetic model. The *elementary level* group contribution rate constants (k_1, k_4, k_5 & k_8) are still defined in the same way as explained in *hypothesis 2*; however, they are now extended to the *overall level* of the reaction, as illustrated in Figure 4-5.

Combination of Overall and Elementary Level Models

$$R_1 + R_2 \underset{k_r}{\overset{k_f}{\rightleftharpoons}} P_1 + P_2$$

$$k_f = (k_1 + k_5) \text{ and } k_r = (k_4 + k_8)$$

$$K_e = \frac{k_f}{k_r}$$

Example: Reaction 2 (Table C-1)

$$k_f = (k_1 + k_5) \text{ and } k_r = (k_4 + k_8)$$

$$\Rightarrow k_1 \leftarrow CH - C_4H_9 \text{ and } k_5 \leftarrow CH - C_3H_7$$

$$\Rightarrow k_4 \leftarrow CH - C_4H_9 \text{ and } k_8 \leftarrow CH - C_3H_7$$

$$K_e = \frac{k_f}{k_r} = \frac{k_{CH-C_4H_9} + k_{CH-C_3H_7}}{k_{CH-C_4H_9} + k_{CH-C_3H_7}} = 1$$

Figure 4-5. Combination of Overall and Elementary Metathesis Models

According to Section 2.8 and 2.9, the rate constants describing a particular direction of the reaction can be additively combined into a single term, which gave rise to the result in Figure 4-5. However, the problem from *hypothesis 1*, that was originally present with modeling at the *overall level* again presents itself (see example Reaction 2 in Figure 4-5), i.e., the equilibrium constant theoretically equating to '1' due to *definition*. The fact that this still occurs indicates a fundamental issue with the way in which the rate constants are defined.

Hypothesis 2 provides the key to solving this problem. Consider the following observation made in the previous section,

$$k_1 = f(M_1, R_1) \text{ \& } k_4 = f(M_2, P_1) \text{ and } k_5 = f(M_2, R_2) \text{ \& } k_8 = f(M_1, P_2)$$

From *hypothesis 2*, it was clear that the *definition* of the rate constants was influenced by the *Alkylidene's* existing group. However, the possibility that the type of *Alkylidene* could have an impact on the *value* of the rate constant was indirectly considered. This is the basis for *hypothesis 3*, in which the type of *Alkylidene* group, that participates in transition state formation, is hypothesized to have a direct impact on the *value* of a defined group contribution rate constant. The difference here is that the elucidation is not made because of the lumping of the alkylidene concentration (indirect); instead, there is a (direct) postulation that the group contribution rate

constant value must be influenced by the type of alkylidene even though no explicit terms indicate this in the rate expression.

Therefore, a modified form of Figure 4-5, incorporating the impact of *hypothesis 3*, is shown in Figure 4-6. It is clear that a reaction, similar to Reaction 2, will *no longer* have an equilibrium constant theoretically equating to '1' as a result of group contribution *definition*.

Modified Combination of Overall and Elementary Level Models

$$R_1 + R_2 \underset{k_r}{\overset{k_f}{\rightleftharpoons}} P_1 + P_2$$

$$k_f = \left(k_{1,M_1} + k_{5,M_2}\right) \text{ and } k_r = \left(k_{4,M_2} + k_{8,M_1}\right)$$

$$K_e = \frac{k_f}{k_r}$$

Example: Reaction 2 (Table C-1)

$$k_f = \left(k_{1,M_1} + k_{5,M_2}\right) \text{ and } k_r = \left(k_{4,M_2} + k_{8,M_1}\right)$$

$$M_1 \leftarrow CH - CH_3 \text{ and } M_2 \leftarrow CH_2$$

$$\Rightarrow k_1 \leftarrow CH - C_4H_9 \text{ and } k_5 \leftarrow CH - C_3H_7$$

$$\Rightarrow k_4 \leftarrow CH - C_4H_9 \text{ and } k_8 \leftarrow CH - C_3H_7$$

$$K_e = \frac{k_f}{k_r} = \frac{k_{CH-C_4H_9,M_1} + k_{CH-C_3H_7,M_2}}{k_{CH-C_4H_9,M_2} + k_{CH-C_3H_7,M_1}} \neq 1$$

Figure 4-6. Modified Combination of Overall and Elementary Metathesis Models

Hypothesis 3 provides a useful simplification of the method proposed in *hypothesis 2*, as it no-longer involves the complexities of transition state dissociation and probabilities. This hypothesis will provide an insight into the type of kinetic behaviour present in this system and will illustrate whether a simplified power-law model can adequately define the chemical interactions occurring in this system.

5. Hypothesis Testing & Rate Parameter Optimization

5.1. Background

The focus of this chapter is to implement the developed framework to determine if a group contribution approach is appropriate for modelling the metathesis process. The section proceeds by implementing a simulation of the reacting system. This will then be used to evaluate the performance of *hypothesis 2 and 3* using sample kinetic parameter sets. Considering that *hypothesis 1* was previously discredited, and that these hypotheses are based on theoretical ideas, it becomes important to investigate the numerical stability of the ODE system that each hypothesis creates, in terms of Section 2.8 and 2.9. Once the stability of each hypothesis is confirmed, the existence of a group contribution modeling approach can be confirmed through multiparameter optimization of the system (identified rate parameters).

5.2. Generalized Model Implementation

The reaction data from Table C-6 was encoded as illustrated in Table D-1, for ease of implementation. The definitions of k_1, k_4, k_5, & k_8 were based on the work done in the Section 4 – These were encoded using the variables shown in Table 5-1.

Table 5-1. Defined Reacting Group Encoding

Groups	**Encoding**
CH-C_4H_9 (largest)	A
CH-C_3H_7	B
CH-C_2H_5	C
CH-CH_3	D
CH_2 (smallest)	E

Further to this, the reacting components were encoded as shown in Table 5-2. These encodings facilitated the implementation of the model in simulation and forms the basis for the investigating the validity of the group contribution approach.

Table 5-2. Chemical Species Encoding

Chemical Species	**Notation**
1-Hexene	1
2-Hexene	2
3-Hexene	3
1-Ethene	4
5-Decene	5
1-Pentene	6
2-Heptene	7
1-Propene	8
4-Nonene	9
1-Butene	10
3-Octene	11
2-Butene	12
4-Octene	13
2-Pentene	14
3-Heptene	15

The lumped rate parameters, as developed in Appendix C5-2, are the combination of either of k_1 and k_8, with M_1, or k_5 and k_4, with M_2, which gives rise to 25 unique rate parameters. Sample values for these parameters are shown in Table D-2. For ease of reference in simulation, these parameters were renamed from k_1 to k_{25} as shown in Table D-2 and cannot be used interchangeably with the previously defined k_1, k_8, k_5 & k_4, as part of the modelling methodology in Section 4.

Utilizing these encodings, with all the developed models and assumptions in this work, the reaction model was implemented as shown in Figure D-1, with the sample code script in Appendix E2. The sample parameter sets shown in Table D-2 were used to perform preliminary tests on the models created using hypotheses 2 & 3, and the results of these investigations are explored in the next section.

5.3. Hypothesis 2 & 3 Stability Evaluation

The main hypothesis was determined using randomly selected parameters in various orders of magnitude – The was aimed at determining the theoretical feasibility of the hypotheses, i.e., hypothesis 2 or 3, and the models based on these. Table D-2 lists the randomly selected values for the defined rate parameters – Larger groups were generally assigned larger defined rate constant values. These values explore different ranges, which should create results that will indicate the stability and viability of each hypothesis.

Table D-3 illustrates the results for hypothesis 2 (model 2), and hypothesis 3 (model 3), when the defined rate parameters hold the values indicated by the respective sets in Table D-2. These results were also graphically displayed in Figures D-4, D-5, and D-6 for Model 2, and Figures D-7, D-8, D-9 for Model 3 (only positive values illustrated). An initial observation is that there are very different flow predictions, for the same parameter set, depending on the model that is used – Model 2 results in positive flow predictions; whereas, Model 3 results in mainly negative flow predictions. This is more clearly seen when comparing Figures D-4 and D-7 for Parameter Set 1, Figures D-5 and D-8 for Parameter Set 2, and Figures D-6 and D-9 for Parameter Set 3 – The reason why so few points exist on the Model 3 plots is due to the fact that only the positive flow predictions are displayed, and, majority of the flows predicted using Model 3 were negative. In the case of Parameter Set 2, Model 3 was incapable of producing a mathematically stable output – Composition profiles peaked to extremely large positive or negative values then instantaneously moved towards a flow prediction of zero; therefore, no flow predictions are visible with Parameter Set 2. Parameter Set 4 predictions are not graphically displayed as the same overall trend is indicated; however, the data is available in Table D-3 for comparison.

These observations indicate that there is inherent instability in Model 3; therefore, the Hypothesis backing it must also lack theoretical soundness. For the reasons described, and the inherent instability observed, Model 3 (Hypothesis 3) will be discarded, and the remainder of this work will focus solely on Model 2 (Hypothesis 2). This also provides preliminary confirmation of the existence of a group contribution approach.

5.4. Single Objective, Multi-parameter Optimization Approaches

Stemming from the development of the model, and simulation, in the previous section, the focus of this work now moves toward optimization of the parameter values, such that the predicted output model flowrates closely match the experimental datasets in Appendix A, for the required conditions. In terms of characterization, the problem that arises is once that involves a single objective (fitness function), with multiple parameters (25 theoretical rate constants) – This is illustrated by Figure 5-1, which is a summary of Figure D-1, and shows a simple schematic of the reaction model.

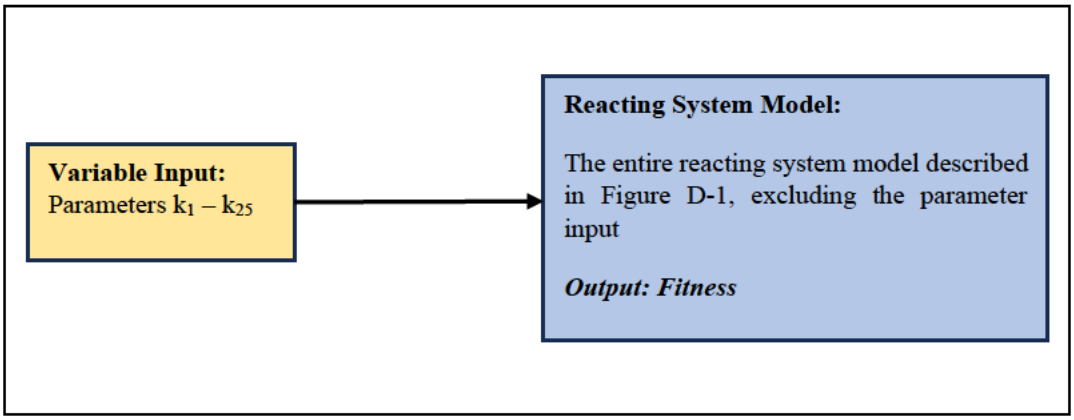


Figure 5-1. Optimization Characteristic

In Figure D-1 and Figure 5-1, the fitness function is referenced by a modified *L2 Norm*. The *L2 Norm* was selected over the *L1 Norm* as it more strongly penalizes deviations in predictions and is preferred for the estimation of parameters (Moravec, 2015). The fitness function was based on a modified *L2 Norm*, and is calculated as shown in Figure 5-2, where the experimental data comes from Appendix A, and the predicted data is the output species flowrates from the reacting system model.

The modified *L2 Norm* was implemented as the magnitude of the usual *L2 Norm* resulted in extremely small fitness values, due to low order or magnitude of the predicted parameters ($\sim 10^{-8}$), which directly impacted the performance of the optimization algorithms. Despite this modification, the overall goal of the optimization process would be to adjust the parameters such that the fitness function value is reduced to zero.

The optimization algorithms that were used for the purpose of adjusting the parameter values to satisfy this objective were the *Particle Swarm Optimisation* (PSO) algorithm, and a *Genetic Algorithm* (GA) method. An overview of the algorithms and an implementation of these with the reacting system model, will be illustrated in the next section. For the optimization process, all experimental carbon chain lengths that exceeded a chain length of 9 were lumped into the ≥C10 grouping.

Experimental Data (E):

Chain Length	*Typical Value*
C1	A
C2	B
C3	C
C4	D
C5	E
C6	F
C7	G
C8	H
C9	I
≥C10 (Lumped)	J

Predicted Data (P):

Chain Length	*Typical Value*
C1	Ap
C2	Bp
C3	Cp
C4	Dp
C5	Ep
C6	Fp
C7	Gp
C8	Hp
C9	Ip
C10	Jp

Fitness Calculation:

$$\sum_{i=1}^{10} \sqrt{(E_i - P_i)^2} \Big/ P_i$$

Figure 5-2. The Fitness Function

5.4.1. Particle Swarm Optimization (PSO)

The *particle swarm optimization* (PSO) technique uses swarm intelligence to optimize a system – Effectively, it considers the impact and randomness of a single 'participant' with reference to the global trend. This involves using a meta-heuristic approach to treat the system as a *black box model*, with as few assumptions as possible, and is useful for the cases where the investigated system consists of a large number of parameters (Mirjalili *et al.*, 2019).

The algorithm, and hyperparameters are further documented in Section 2.15 of the literature review; however, the focus here is implementation and use. The detailed implementation of the algorithm, with reference to the generalized model in Figure D-1, is shown in Figure D-2.

The set-up illustrated in Figure D-2, uses a manually inputted seed set of parameters to initialize the process, which is an important factor in the methodology used to optimize the system parameters. *Python* code scripts illustrating the implementation of the optimization technique are provided in Appendix E3.

5.4.2. Genetic Algorithms (GAs)

Genetic algorithms (GAs) also form part of population-based metaheuristic techniques and are used to optimize systems with a large number of parameters (Mirjalili *et al.*, 2019). GAs can be implemented in many ways and use evolutionary techniques to introduce randomness into the search. The focus in this implementation was for real coded GAs, given the low order of magnitude of the parameter values (Katoch, Chauhan, and Kumar, 2020).

Tournament selection, *SBX crossover*, and *Polynomial mutation* were the main types of steps used in implementation. The algorithm and hyperparameters are further explored in in Section 2.16 of the literature review; however, the focus here is implementation and use. A flowchart detailing the implementation of the algorithm, with reference to the generalized model in Figure D-1, is shown in Figure D-3.

Similar to the PSO set-up, the set-up illustrated in Figure D-3 uses a manually inputted seed set of parameters to initialize the process. *Python* code scripts illustrating the implementation of the optimization technique are provided in Appendix E4.

5.5. Regression Methodology

The methodology used to optimize the system parameters were based on the deficiencies observed during testing of the PSO (Figure D-2) and GA (Figure D-3) optimization set-up. Despite adjusting the relevant hyperparameters, it was found that the PSO algorithm performed better at searching globally, and the GA performed better when searching locally.

The method used to conduct the regressions involves exploiting this observation, which involved initially searching a wider space using the PSO technique with a seed parameter set (all parameter values set to 1×10^{-6}), and then repetitively using the outputs as seed values for either the GA or PSO techniques, in a more localized space. This process is illustrated in Figure 5-3.

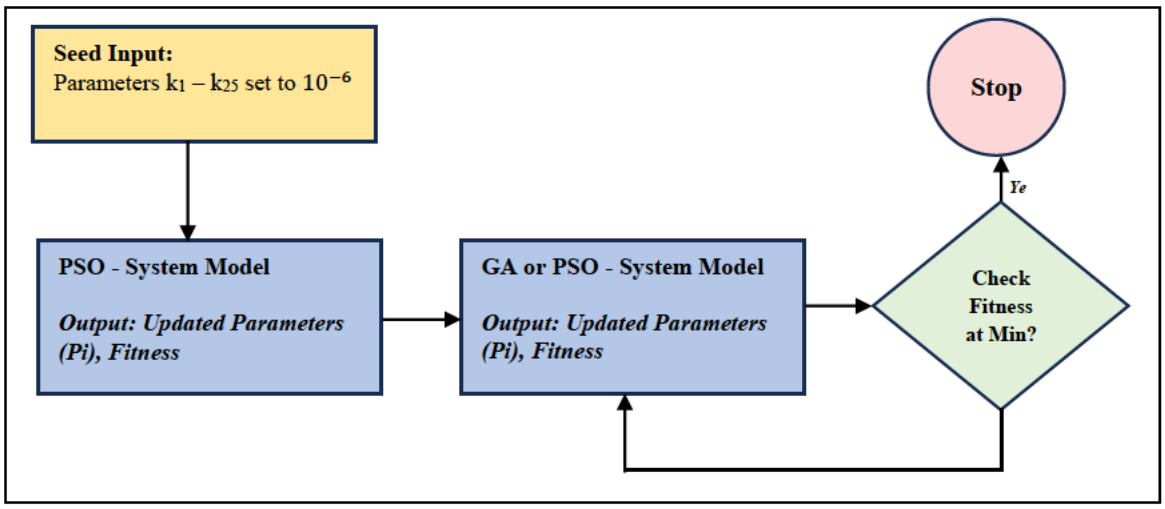


Figure 5-3. Regression Methodology Flowchart

The identified rate parameters resulting from the process illustrated in Figure 5.3 are provided in Table D-4, of Appendix D3.

5.6. Regression Results & Kinetic Parameters

The identified kinetic parameters were determined using the literature experimental data in Appendix A, available at a range of temperatures (420°C - 460°C). Given that the system consisted of 25 parameters that interact within the reaction model, it becomes inherently difficult to regress parameters toward fixed values, as multiple combinations of values for different parameters will yield a similar effect. At a rudimentary level, this effect was captured and accounted for by using multiple sets of experimental data, at a range of temperatures.

The *Arrhenius plots* for each rate parameter are shown in Figures 5-4 to 5-28 – These were constructed using the data in Table D-4, and the method in Section 2.13. A summary of the activation energy (E_a), and pre-exponential factor (A_o), obtained using the Arrhenius plots are provided in Table 5-3.

Table 5-3. Identified Arrhenius Constants

Parameter $\left[\frac{m^3}{sec.kg}\right]$	Attaching Group	Existing Group	E_a [J/mol]	A_o $\left[\frac{m^3}{sec.kg}\right]$	P_F	$P_{intercept}$	P_{slope}	R^2 Bound	R^2
k1	A	A	1,826E+05	9,005E+09	0,0053	0,0226	0,0053	>0.4	0,6420
k2	A	B	5,053E+05	3,625E+31	0,2290	0,2994	0,2290	<0.4	0,1750
k3	A	C	6,404E+05	1,884E+42	0,0358	0,0529	0,0358	>0.4	0,4425
k4	A	D	1,411E+04	2,166E-04	0,9739	0,9076	0,9739	<0.4	0,0001
k5	A	E	2,665E+04	4,357E-02	0,6068	0,7188	0,6068	<0.4	0,0346
k6	B	A	4,258E+05	3,137E+27	0,0639	0,0950	0,0639	>0.3	0,3660
k7	B	B	3,229E+04	1,225E-02	0,9207	0,9359	0,9207	<0.3	0,0013
k8	B	C	1,436E+05	1,021E+07	0,0739	0,2083	0,0739	>0.3	0,3456
k9	B	D	1,417E+05	4,206E+06	0,2331	0,4346	0,2331	<0.3	0,1722
k10	B	E	1,255E+05	7,353E+05	0,0272	0,1237	0,0272	>0.3	0,4764
k11	C	A	2,297E+05	4,711E+12	0,3920	0,5152	0,3920	<0.3	0,0928
k12	C	B	8,135E+04	1,657E+02	0,4521	0,7762	0,4521	<0.3	0,0724
k13	C	C	8,648E+05	1,994E+58	0,0111	0,0166	0,0111	>0.3	0,5739
k14	C	D	-5,436E+05	6,104E-46	0,2731	0,2185	0,2731	<0.3	0,1476
k15	C	E	3,057E+05	2,141E+18	0,3534	0,4436	0,3534	<0.3	0,1082
k16	D	A	5,776E+05	1,211E+38	0,0806	0,1098	0,0806	>0.3	0,3333
k17	D	B	8,531E+05	6,977E+57	0,0155	0,0220	0,0155	>0.3	0,5401
k18	D	C	2,116E+05	5,678E+11	0,0051	0,0200	0,0051	>0.3	0,6461
k19	D	D	5,284E+05	2,390E+34	0,0972	0,1341	0,0972	>0.3	0,3060
k20	D	E	7,071E+05	2,031E+47	0,0624	0,0837	0,0624	>0.3	0,3692
k21	E	A	1,547E+05	6,031E+07	0,0482	0,1488	0,0482	>0.3	0,4041
k22	E	B	1,274E+05	5,784E+05	0,2262	0,4419	0,2262	<0.3	0,1769
k23	E	C	-4,650E+04	4,573E-09	0,8727	0,6964	0,8727	<0.3	0,0034
k24	E	D	1,246E+05	2,153E+05	0,3270	0,5590	0,3270	<0.3	0,1199
k25	E	E	1,221E+02	3,423E+04	0,6873	0,8378	0,6873	<0.3	0,0213

Where, A is *ch-c₄h₉* (largest), B is *ch-c₃h₇*, C is *ch-c₂h₅*, D is *ch-ch₃*, and E is *ch₂* (smallest)

A key literature-based assumption, which was the basis for *hypothesis 2*, is that the size of the attaching group directly influences the rate of the metathesis reaction, by aiding or slowing the transition state formation process (Kapteijn and Mol, 1981). Therefore, the rate of a reaction step should be slower when larger groups are involved – A slower reaction occurs when the activation energy is high, and the pre-exponential factor is low. This can be inferred from the definition of the activation energy (E_a), and pre-exponential factor (A_o), in Section 2.13.

From Section 2.10, 2.11 and 2.13, a higher step activation energy implies that more energy would be required for that phase of the reaction to occur, which would slow down the process relative to a lower activation energy step. A higher pre-exponential factor implies that there would be a greater likelihood of the respective reactants (groups) colliding. The results obtained will be assessed in terms of these principles.

This trend is visible in the data of Table 5-3; however, there are numerous other chemical interaction factors that participate in the transition state dissociation and reformation steps that occur during the metathesis process. Therefore, it makes sense that the uniqueness of each rate parameter was accounted for using the *attaching* and *existing* group definition, of *hypothesis 2*.

Each parameter applies to the various steps of the overall metathesis reaction – The method used to define these parameters was detailed in Section 4, with the numerical quantification of the kinetics of each step being a key process in modeling the kinetics of the metathesis system.

The values illustrated in the *Arrhenius plots* that are constructed in this section can be converted into the values shown in Table 5-3 using the method illustrated in Section 2.13. The P-values describing the fit for all identified parameters are provided in Table 5-3. These P-values were calculated for overall significance (PF), and the model coefficients (intercept and slope).

P-value Analysis – Overall Significance

The F significance (PF) value is the P-value for the F-test, and this indicates whether the independent variables [*1/T*] are globally related to the dependent variable [*ln(k)*] – Since the PF value is below the significance threshold of 0.05 for *k1, k3, k6, k10, k13, k17, k18, k21,* it implies that the respective linear model more effectively captures the dependent variable's variability, when compared to a model with no independent variables. Therefore, the respective models are statistically significant.

Since the PF value is above the significance threshold of 0.05 for *k2, k4, k5, k7, k8, k9, k11, k12, k14, k15, k16, k19, k20, k22, k23, k24, k25,* it implies that the respective model does not capture the dependent variable's variability, when compared to a linear model with no independent variables. Therefore, the respective models are not statistically significant.

P-value Analysis – Intercept Significance

The P-values for the intercept coefficient indicates whether the dependent variable [*ln(k)*] is statistically significant, and these values are illustrated in the $P_{intercept}$ column of Table 5-3. Since the P-value is below the significance threshold of 0.05 for the intercept of *k1, k13, k17, k18*, the null hypothesis that the coefficient equates zero can be rejected. This implies that a relationship between the dependent variable and intercept coefficient exists. Therefore, the respective pre-exponential factors determined from this process are likely to be good representations of reality.

Since the P-value is above the significance threshold of 0.05 for the intercept of *k2, k3, k4, k5, k6, k7, k8, k9, k10, k11, k12, k14, k15, k16, k19, k20, k21, k22, k23, k24, k25*, the null hypothesis that the coefficient equates to zero cannot be rejected with certainty. This implies that a relationship between the dependent variable and the intercept coefficient may not exist. Therefore, the respective pre-exponential factors determined from this process are likely to be poor representations of reality.

P-value Analysis – Slope Significance

The P-values for the slope coefficient indicates whether the dependent variable [*ln(k)*] is statistically significant, and these values are illustrated in the P_{slope} column of Table 5-3. Since the P-value is below the significance threshold of 0.05 for the slope of *k1, k3, k10, k13, k17, k18, k21,* the null hypothesis that the coefficient equates zero can be rejected. This implies that a relationship between the dependent variable and slope coefficient exists. Therefore, the respective activation energy values determined from this process are likely to be good representations of reality.

Since the P-value is above the significance threshold of 0.05 for the slope of *k2, k4, k5, k6, k7, k8, k9, k11, k12, k14, k15, k16, k19, k20, k22, k23, k24, k25,* the null hypothesis that the coefficient equates to zero cannot be rejected with certainty. This implies that a relationship between the dependent variable and the slope coefficient may not exist. Therefore, the activation energy determined from this process is likely to be a poor representation of reality.

Arrhenius plots - k1, k2, k3, k4 & k5

Significant spread exists in parameter *k1* when fitted for the lower temperature experimental datasets. This is illustrated in Figure 5-4, which indicates a coefficient of determination (R^2) value of 0.642 – This indicates that variance in the actual data is accounted for by the linear model. Therefore, the values shown in Table 5-3 would be more likely to be a realistic representation as the variance across all the investigated temperatures are suitably accounted for.

This implies that the constants in Table 5-3, determined for group A attaching to group A, would likely be a good representation of reality. Therefore, this will be used as a benchmark to verify the applicability of the size-based effect of the participating groups, where the attaching group is A (parameter *k1* to *k5*). This will only be considered where R^2 values are above the selected relative threshold of 0.4. Considering that the values observed for parameter *k1* are applicable to the largest groups, the underpinning assumption requires that all the calculated constants for the applicable parameters be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k1*.

For the regression analysis performed in Figure 5-4 to 5-8, the detailed statistical parameters are provided in Table 5-3. Considering that this is based on the *Arrhenius transformation* of Section 2.13, the independent variable in the regression would be *1/T*, and the dependent variable would be *ln(k)*.

Similarly, R^2 values were 0.1750, 0.4425, 0.0001, and 0.0346 for parameters *k2* in Figure 5-5, *k3* in Figure 5-6, *k4* in Figure 5-7, and *k5* in Figure 5-8, respectively. In this parameter set, group A is the attaching group, while the existing group is varied, as illustrated in Table 5-3. From the selected threshold of 0.4, the linear models for parameters *k2*, *k4* and *k5*, do not adequately capture the variance in the actual data since it is below the selected relative threshold; therefore, the identified rate parameters in Table 5-3 will not be a good relative representation of reality. However, the linear model for parameter *k3* adequately captures the variance in the actual data since it is above the selected relative threshold, which implies that the identified rate parameters in Table 5-3 will be a good relative representation of reality. The variability in the fit observed is likely a consequence of the large number of parameters that have been simultaneously fitted.

The values observed for parameter *k2* are applicable to a large (A) and moderately large sized (B) group, parameter *k4* are applicable to a large (A) and a moderately small (D) reacting group, and parameter *k5* are applicable to the largest (A), and smallest (E) reacting group, the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k1*. However, these will not necessarily be observed due to the poor relative fit for these parameters.

The values observed for parameter *k3* are applicable to a large (A) and medium sized (C) group; therefore, the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k1*. This is observed in the constants determined for parameter *k3*; however, the magnitude of the determined values likely hold a substantial degree of error.

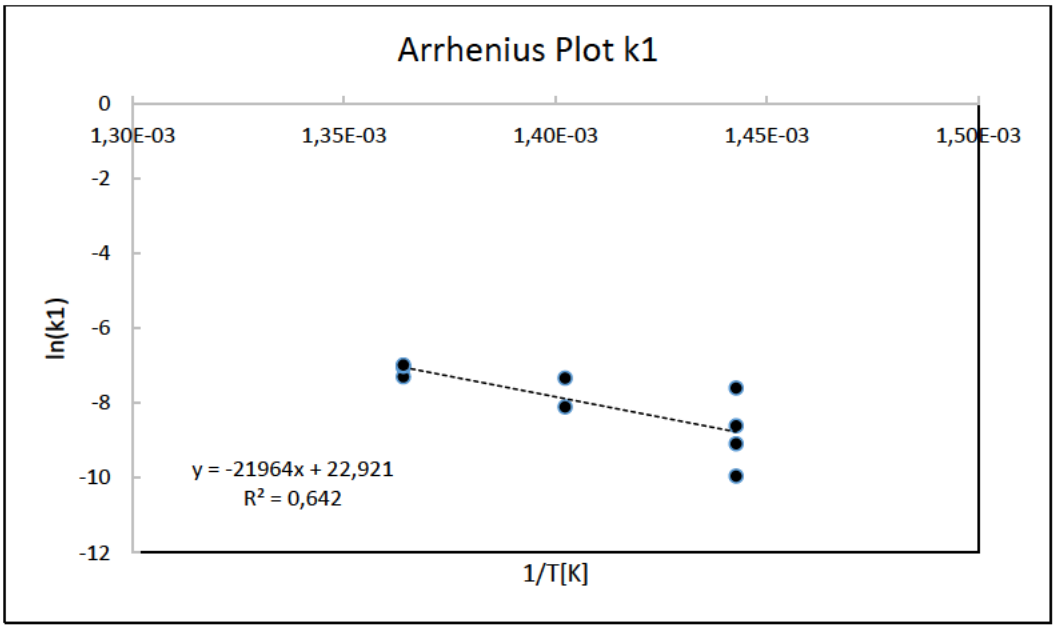


Figure 5-4. Arrhenius Plot for Parameter k1

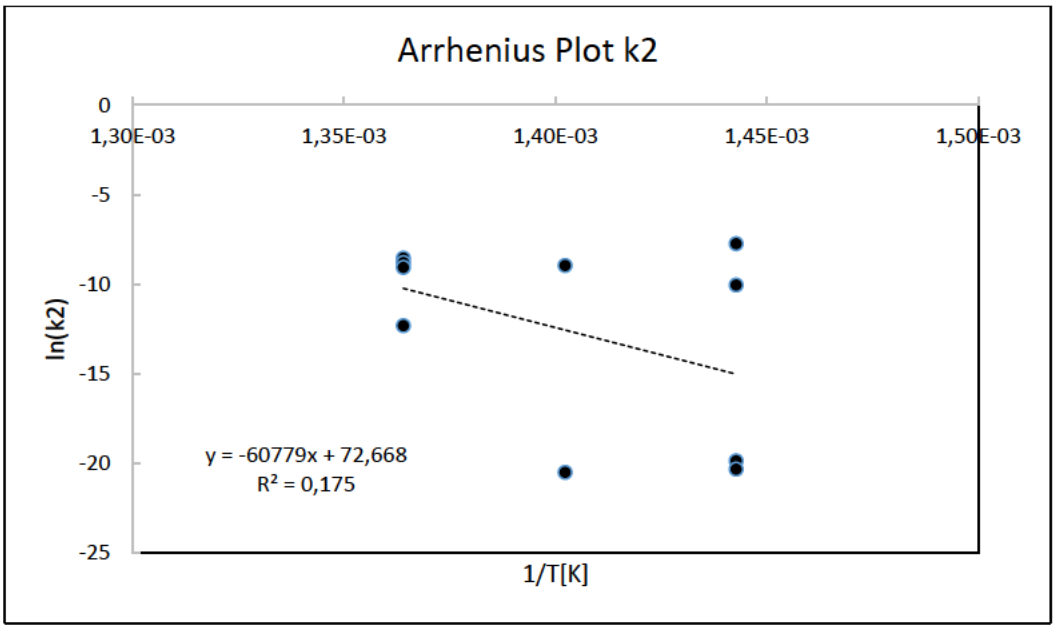


Figure 5-5. Arrhenius Plot for Parameter k2

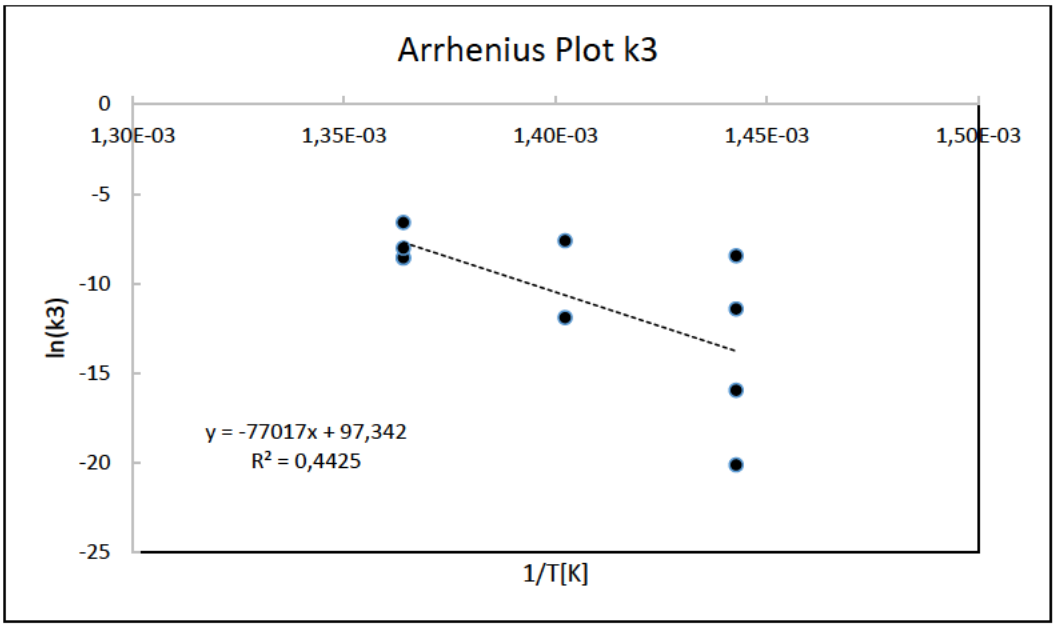


Figure 5-6. Arrhenius Plot for Parameter k3

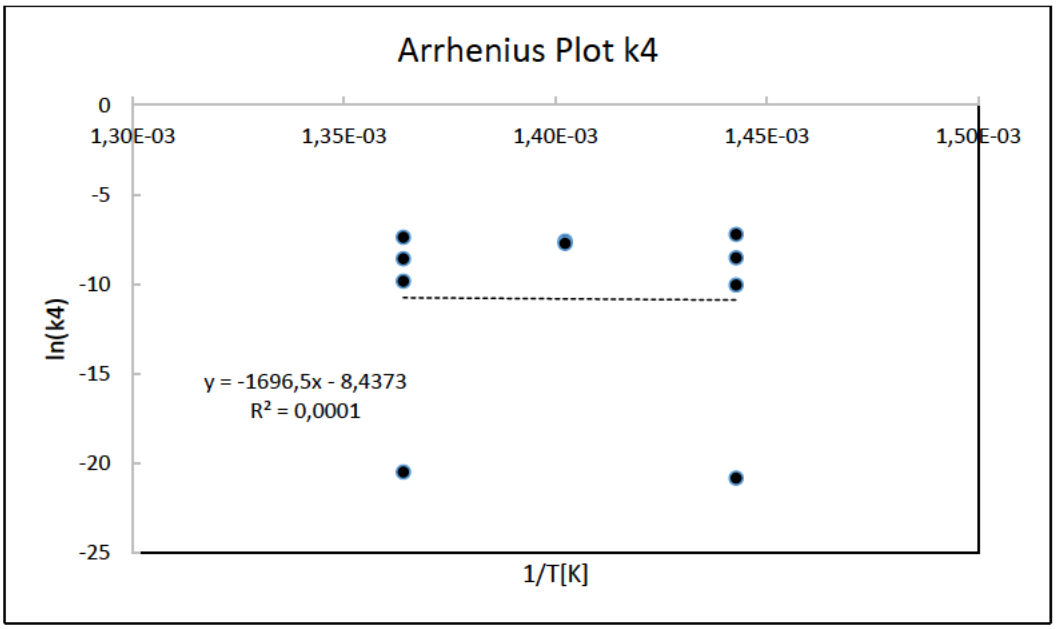


Figure 5-7. Arrhenius Plot for Parameter k4

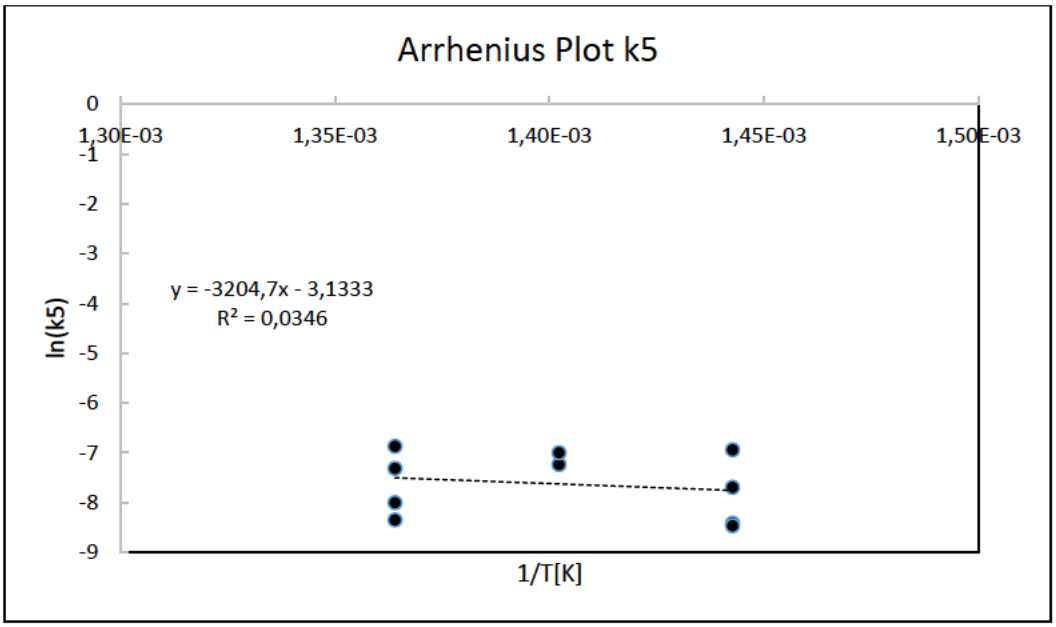


Figure 5-8. Arrhenius Plot for Parameter k5

Arrhenius plots – k6, k7, k8, k9 & k10

Significant spread exists in parameter *k6* when fitted for the lower temperature experimental datasets. This is illustrated in Figure 5-9, which indicates a coefficient of determination (R^2) value of 0.3660 – This indicates that variance in the actual data is partially accounted for by the linear model. Therefore, the values shown in Table 5-3 would be more likely to be a suitable representation of reality, as the variance across all the investigated temperatures are adequately accounted for.

This implies that the constants in Table 5-3, determined for group B attaching to group A, would likely be an adequate representation of reality. Therefore, this will be used as a benchmark to verify the applicability of the size-based effect of the participating groups, where the attaching group is B (parameter *k6* to *k10*). This will only be considered where R^2 values are above the selected relative threshold of 0.3. Considering that the values observed for parameter *k6* are applicable to the larger groups (B & A), the assumption requires that the determined constants for all the applicable parameters, be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k6*.

For the regression analysis performed in Figure 5-9 to 5-13, the detailed statistical parameters are provided in Table 5-3. Considering that this is based on the *Arrhenius transformation* of Section 2.13, the independent variable in the regression would be *1/T*, and the dependent variable would be *ln(k)*.

Similarly, R^2 values were 0.0013, 0.3456, 0.1722, and 0.4764 for parameters *k7* in Figure 5-10, *k8* in Figure 5-11, *k9* in Figure 5-12, and *k10* in Figure 5-13, respectively. In this parameter set, group B is the attaching group, while the existing group is varied, as illustrated in Table 5-3. From the selected threshold of 0.3, the linear models for parameters *k7* and *k9*, do not adequately capture the variance in the actual data since it is below the selected relative threshold; therefore, the identified rate parameters in Table 5-3 will not be a good relative representation of reality. However, the linear model for parameter *k8* and *k10* adequately captures the variance in the actual data since it is above the selected relative threshold, which implies that the identified rate parameters in Table 5-3 will be a good relative representation of reality. The variability in the fit observed is likely a consequence of the large number of parameters that have been simultaneously fitted.

The values observed for parameter *k7* are applicable to two moderately large sized reacting groups (B & B), and parameter *k9* are applicable to a moderately large (B) and moderately small (D) sized reacting group, the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k6*. However, these will not necessarily be observed due to the poor relative fit for these parameters.

The values observed for parameter *k8* are applicable to a moderately large (B) and medium sized (C) group, and parameter *k10* are applicable to a moderately large group (B) and the smallest group (E); therefore, the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when

compared to the values determined for parameter *k6*. This is observed only in the activation energy for parameter *k8* and *k10*; however, the magnitude of the determined values likely hold a substantial degree of error.

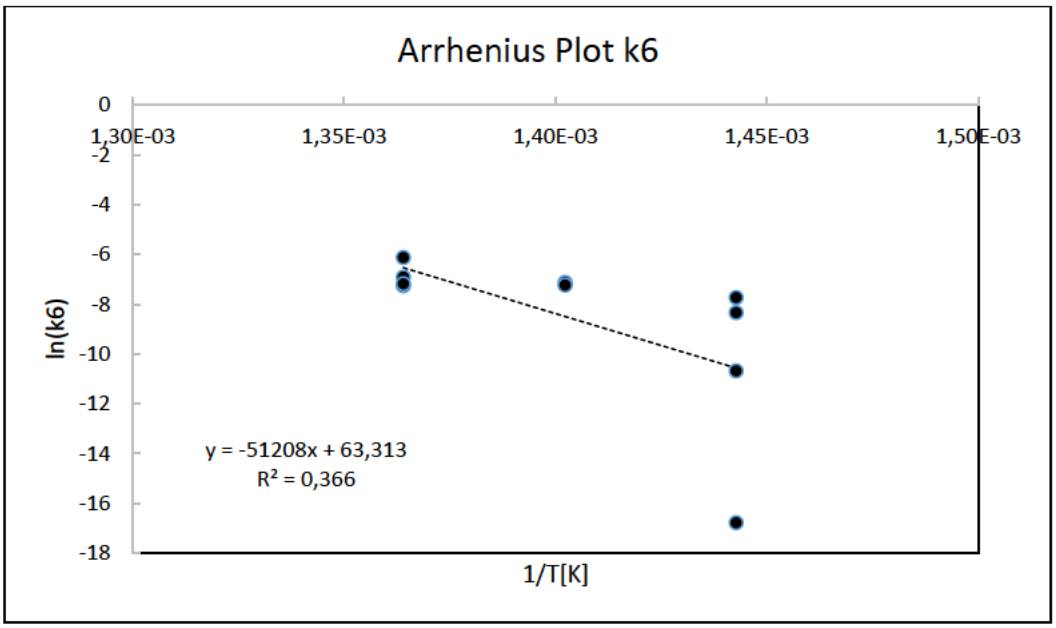


Figure 5-9. Arrhenius Plot for Parameter k6

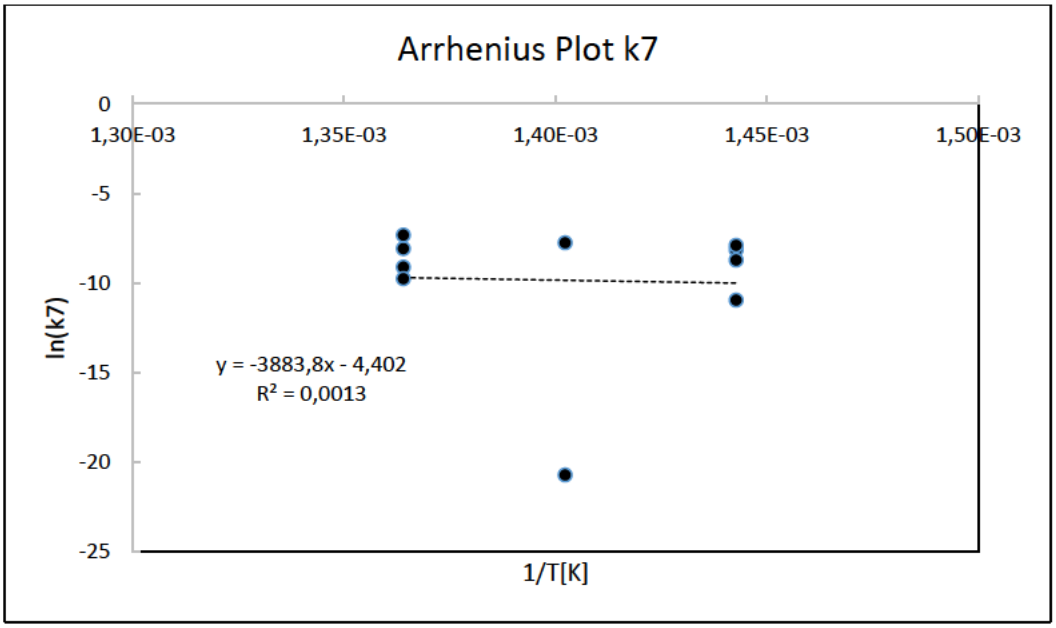


Figure 5-10. Arrhenius Plot for Parameter k7

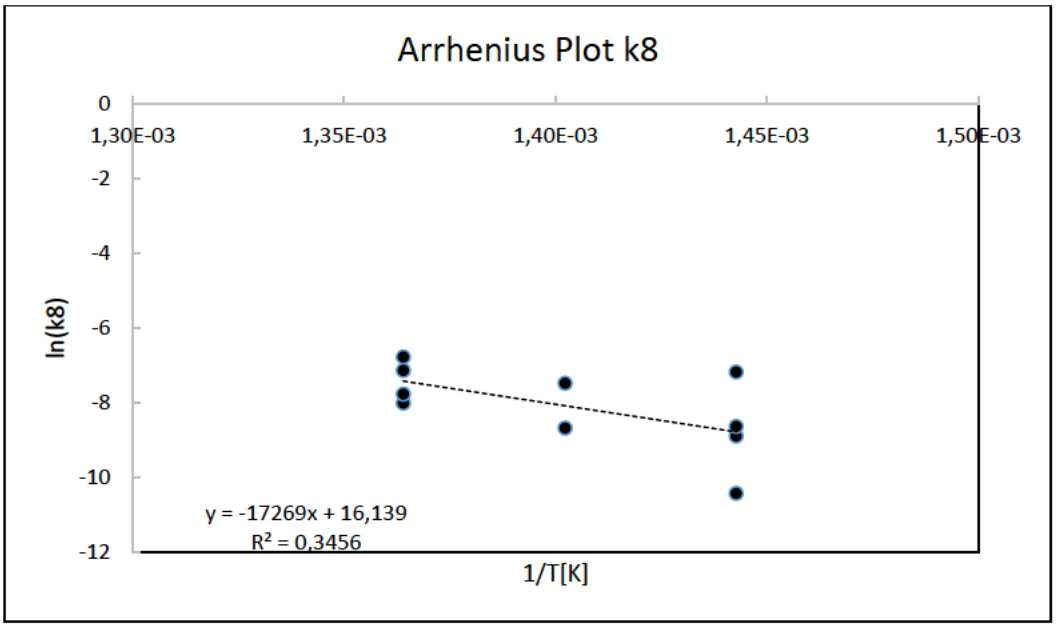


Figure 5-11. Arrhenius Plot for Parameter k8

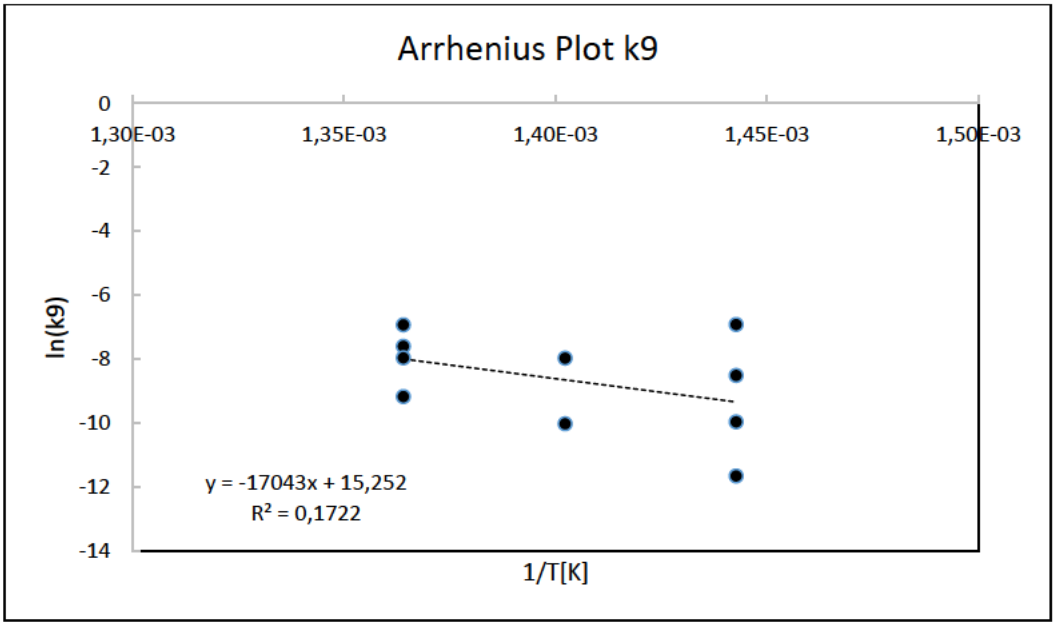


Figure 5-12. Arrhenius Plot for Parameter k9

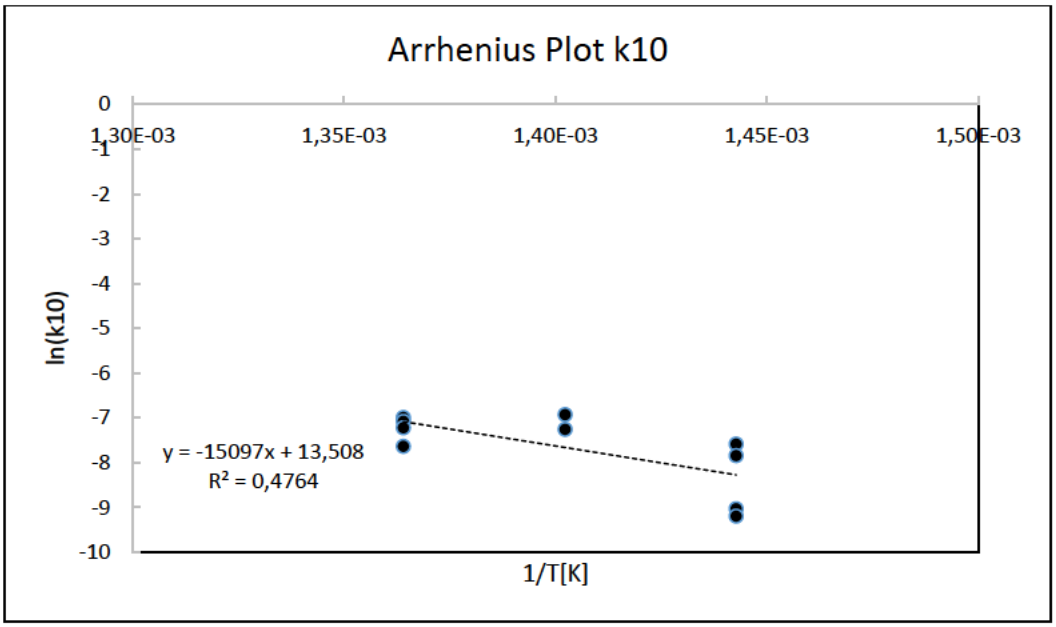


Figure 5-13. Arrhenius Plot for Parameter k10

Arrhenius plots – k11, k12, k13, k14 & k15

Significant spread exists in parameter *k11* when fitted for the experimental datasets. This is illustrated in Figure 5-14, which indicates a coefficient of determination (R^2) value of 0.0928 – This indicates that variance in the actual data is not accounted for by the linear model. Therefore, the values shown in Table 5-3 would be less likely to be a suitable representation of reality, as the variance across all the investigated temperatures are not adequately accounted for.

This implies that the constants in Table 5-3, determined for group C attaching to group A, would likely be a poor representation of reality. This will be used as a benchmark to verify the applicability of the size-based effect of the participating groups, where the attaching group is C (parameter *k11* to *k15*), due to the size of the groups involved. This will only be considered where R^2 values are above the selected relative threshold of 0.3. Considering that the values observed for parameter *k11* are applicable to a medium sized (C), and the largest (A) group, the assumption requires that all the determined constants for the applicable parameters be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k11*.

For the regression analysis performed in Figure 5-14 to 5-18, the detailed statistical parameters are provided in Table 5-3. Considering that this is based on the *Arrhenius transformation* of Section 2.13, the independent variable in the regression would be *1/T*, and the dependent variable would be *ln(k)*.

Similarly, R^2 values were 0.0724, 0.5739, 0.1476, and 0.1082 for parameters *k12* in Figure 5-15, *k13* in Figure 5-16, *k14* in Figure 5-17, and *k15* in Figure 5-18, respectively. In this parameter set, group C is the attaching group, while the existing group is varied, as illustrated in Table 5-3. From the selected relative threshold of 0.3, the linear models for parameters *k12*, *k14,* and *k15*, do not adequately capture the variance in the actual data since it is below the selected relative threshold; therefore, the identified rate parameters in Table 5-3 will not be a good relative representation of reality. However, the linear model for parameter *k13* adequately captures the variance in the actual data since it is above the selected relative threshold, which implies that the identified rate parameters in Table 5-3 will be a good relative representation of reality. The variability in the fit observed is likely a consequence of the large number of parameters that have been simultaneously fitted.

The values observed for parameter *k12* are applicable to a medium sized (C) and a large (A) group, parameter *k14* are applicable to a medium sized (C) and a moderately small (D) group, and parameter *k15* are applicable to a medium sized (C) and the smallest group (E), the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k11*. However, these will not necessarily be observed due to the poor relative fit for these parameters.

Despite the potential inaccuracies, the activation energy value changes from a positive (endothermic) to a negative (exothermic) value, when *k14* is compared to *k11*. This indicates that there is a potential chemical effect

that occurs when the groups defining *k14* interact and illustrates that there is likely more than steric factors that can hinder or promote the reaction.

The values observed for parameter *k13* are applicable to two moderately sized (C & C); therefore, the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k11*. This is observed only in the pre-exponential factor for parameter *k13*; however, the magnitude of the determined values likely hold a substantial degree of error.

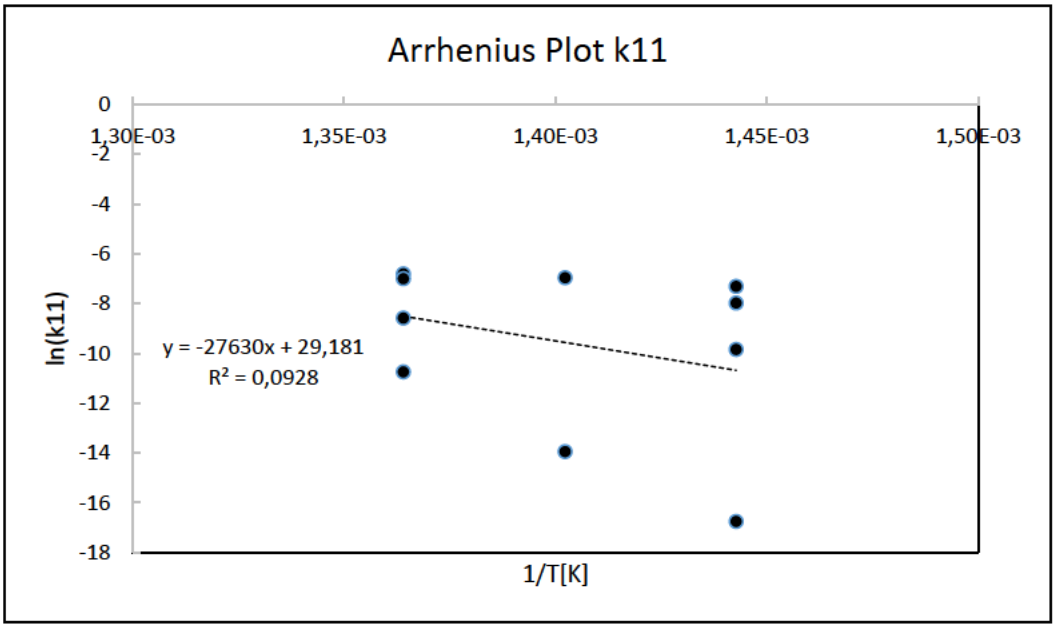


Figure 5-14. Arrhenius Plot for Parameter k11

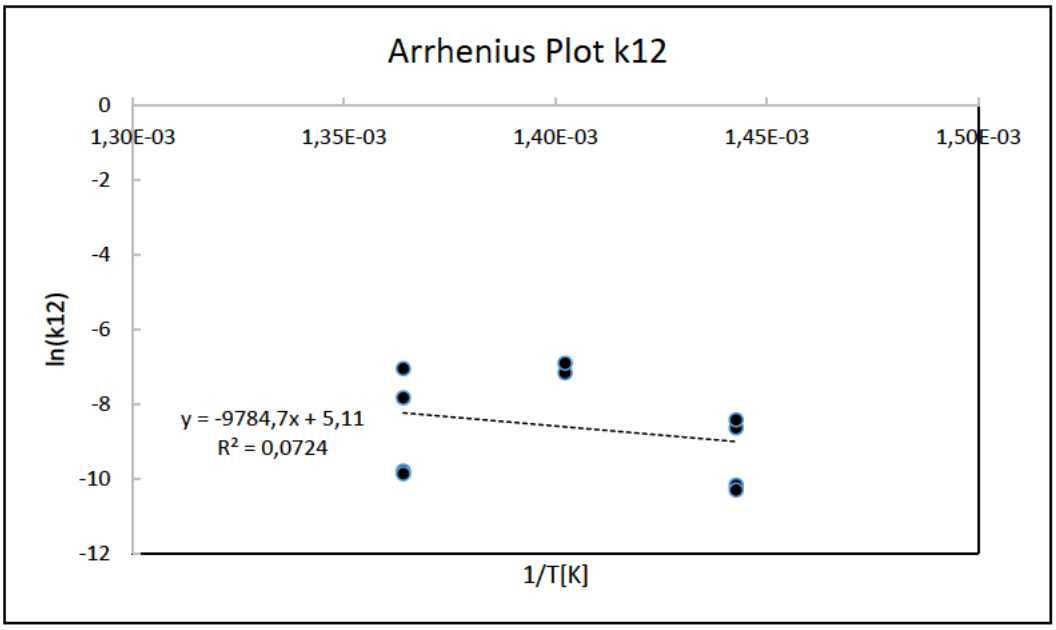


Figure 5-15. Arrhenius Plot for Parameter k12

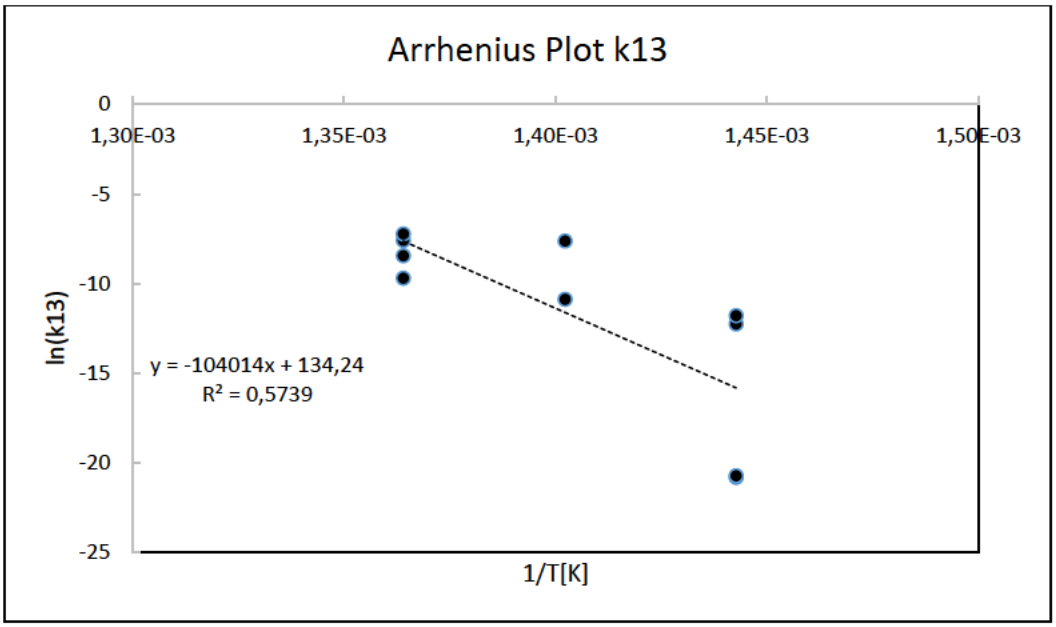


Figure 5-16. Arrhenius Plot for Parameter k13

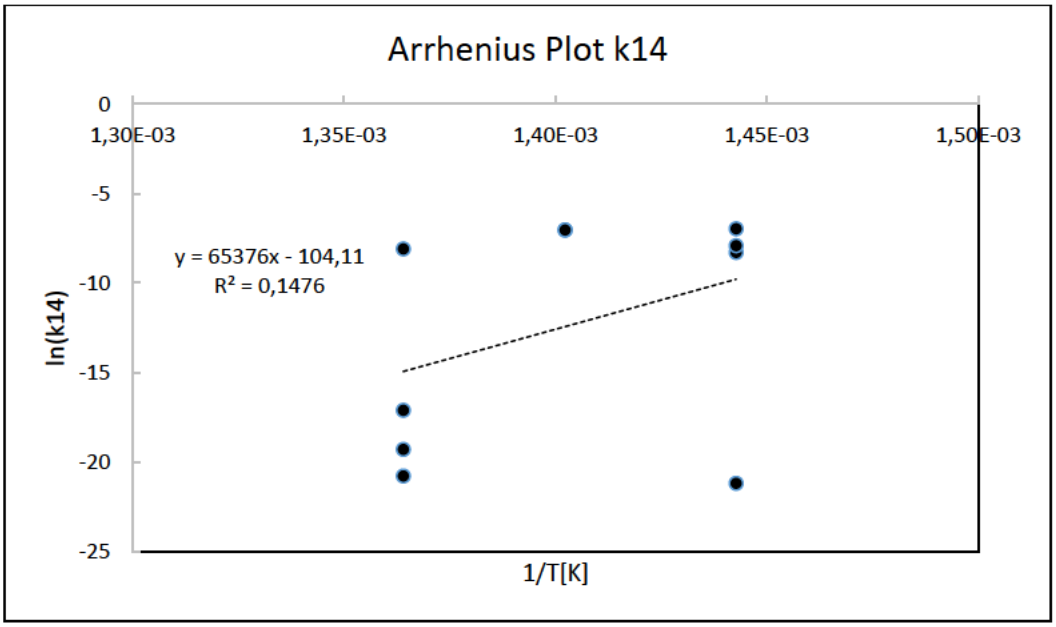


Figure 5-17. Arrhenius Plot for Parameter k14

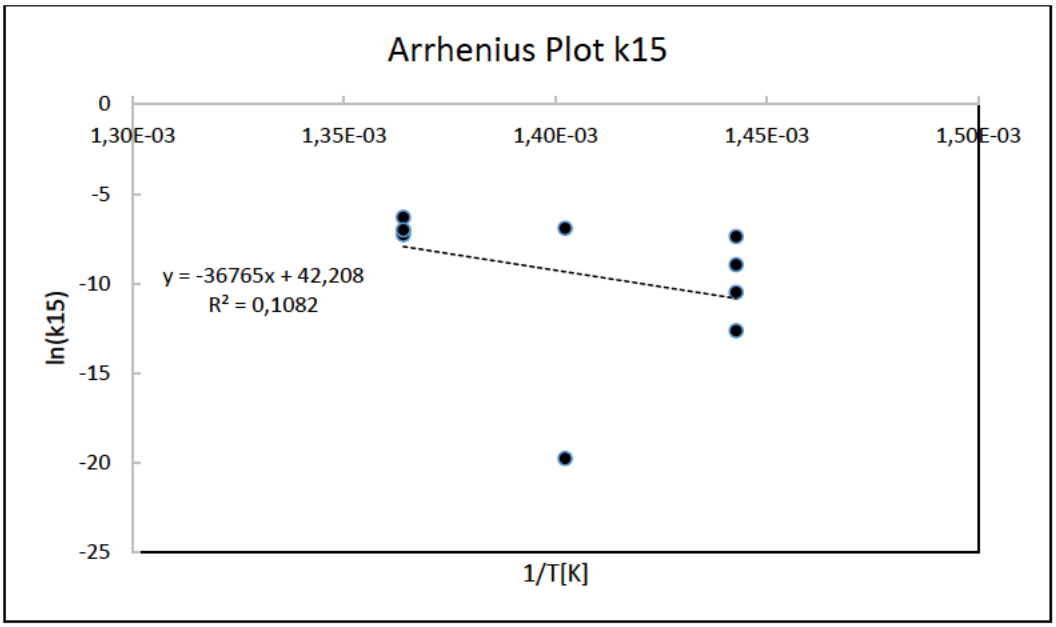


Figure 5-18. Arrhenius Plot for Parameter k15

Arrhenius plots – k16, k17, k18, k19 & k20

Significant spread exists in parameter *k16* when fitted for the experimental datasets. This is illustrated in Figure 5-19, which indicates a coefficient of determination (R^2) value of 0.3333 – This indicates that variance in the actual data is partially accounted for by the linear model, since it is above the selected threshold value. Therefore, the values shown in Table 5-3 would be more likely to be a suitable representation of reality, as the variance across all the investigated temperatures are adequately accounted for.

This implies that the constants in Table 5-3, determined for group D attaching to group A, would likely be an adequate representation of reality. Therefore, this will be used as a benchmark to verify the applicability of the size-based effect of the participating groups, where the attaching group is D (parameter *k16* to *k20*). This will only be considered where R^2 values are above the selected relative threshold of 0.3. Considering that the values observed for parameter *k16* are applicable to a moderately small (D) and the largest (A) group, the assumption requires that all the calculated constants for the applicable parameters should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k16*.

For the regression analysis performed in Figure 5-19 to 5-23, the detailed statistical parameters are provided in Table 5-3. Considering that this is based on the *Arrhenius transformation* of Section 2.13, the independent variable in the regression would be *1/T*, and the dependent variable would be *ln(k)*.

Similarly, R^2 values were 0.5401, 0.6461, 0.3060, and 0.3692 for parameters *k17* in Figure 5-20, *k18* in Figure 5-21, *k19* in Figure 5-22, and *k20* in Figure 5-23, respectively. In this parameter set, group D is the attaching group, while the existing group is varied, as illustrated in Table 5-3. The linear model for parameters *k17*, *k18, k19,* and *k20* adequately captures the variance in the actual data since it is above the selected relative threshold, which implies that the identified rate parameters in Table 5-3 will be a good representation of reality. The variability in the fit observed is likely a consequence of the large number of parameters that have been simultaneously fitted.

The values observed for parameter *k17* are applicable to a moderately small (D) group and a moderately large group (B), parameter *k18* are applicable to a moderately small group (D) and a medium sized (C) group, parameter *k19* are applicable to two moderately small groups (D & D), and parameter *k20* are applicable to a moderately small group (D) and the smallest group (E); therefore, the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k16*. This is observed only in the pre-exponential factor for parameter *k17* and *k20,* and the activation energy for parameter *k18* and *k19*; however, the magnitude of the determined values likely hold a substantial degree of error.

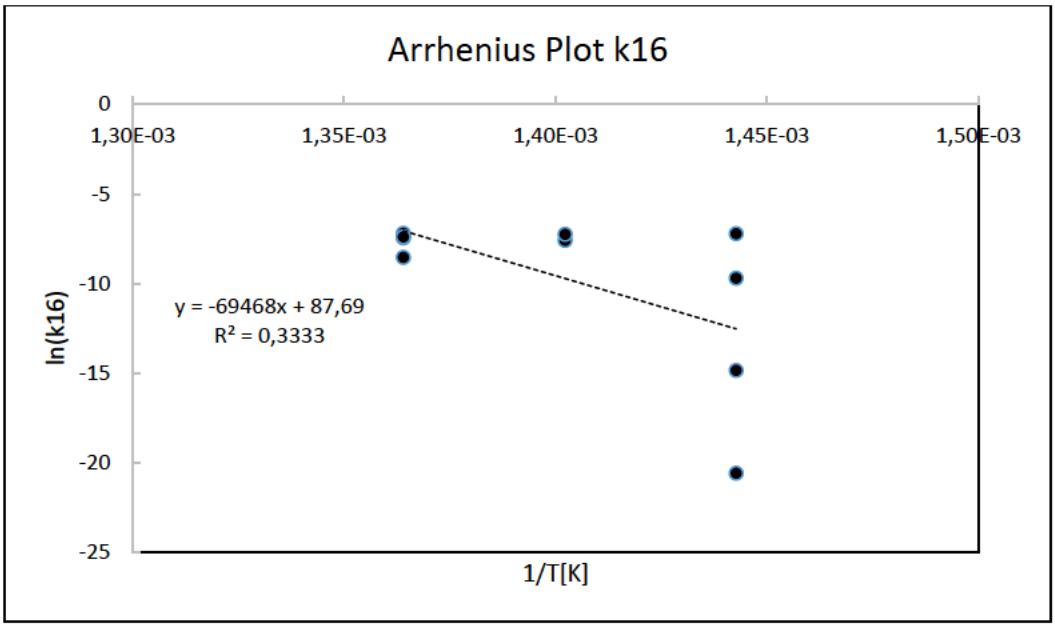


Figure 5-19. Arrhenius Plot for Parameter k16

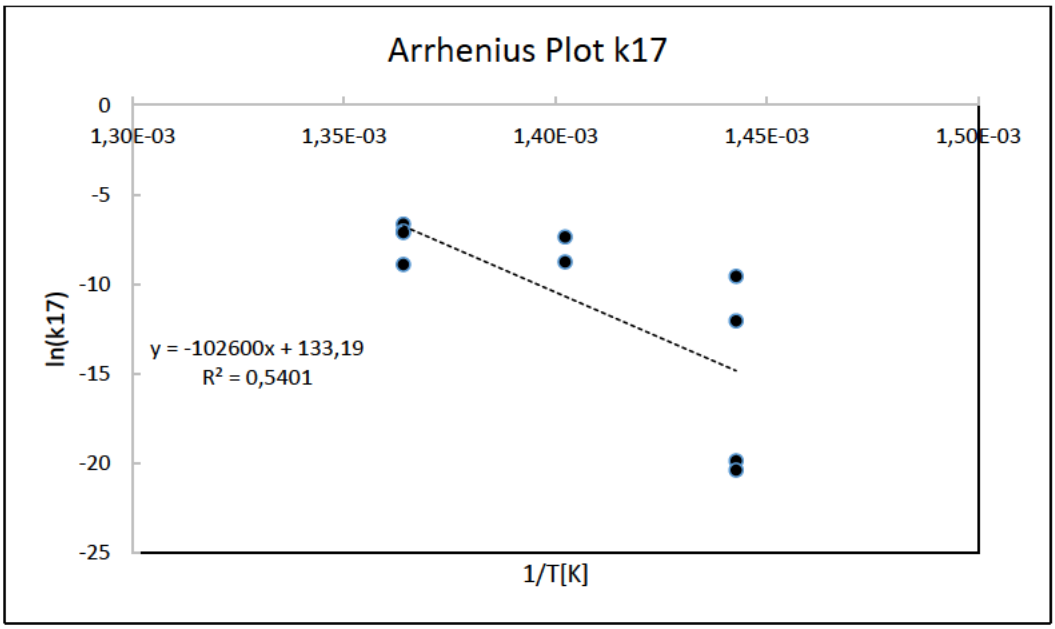


Figure 5-20. Arrhenius Plot for Parameter k17

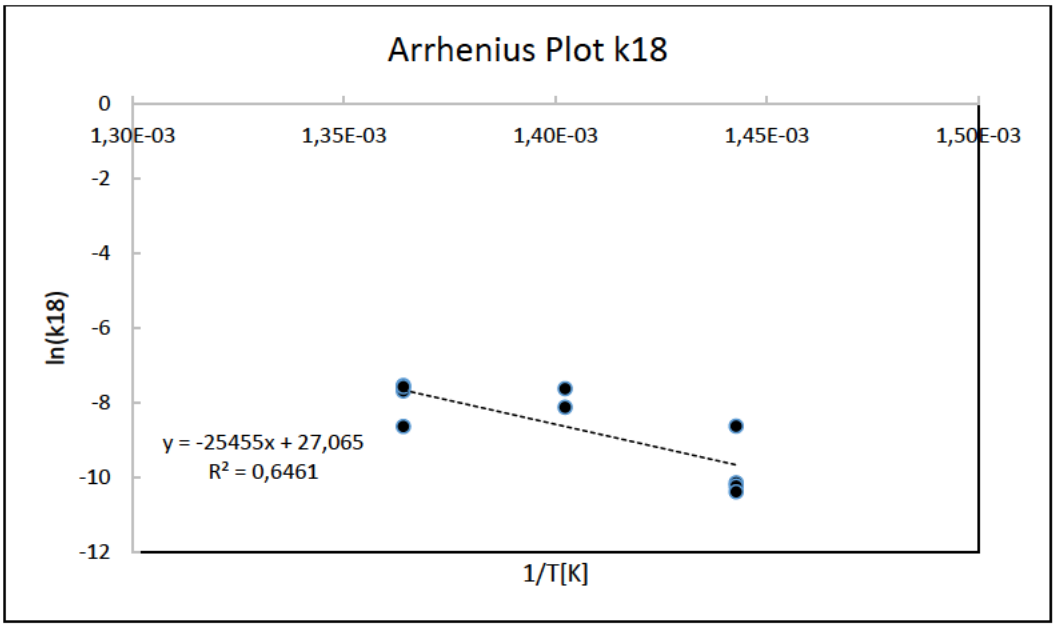


Figure 5-21. Arrhenius Plot for Parameter k18

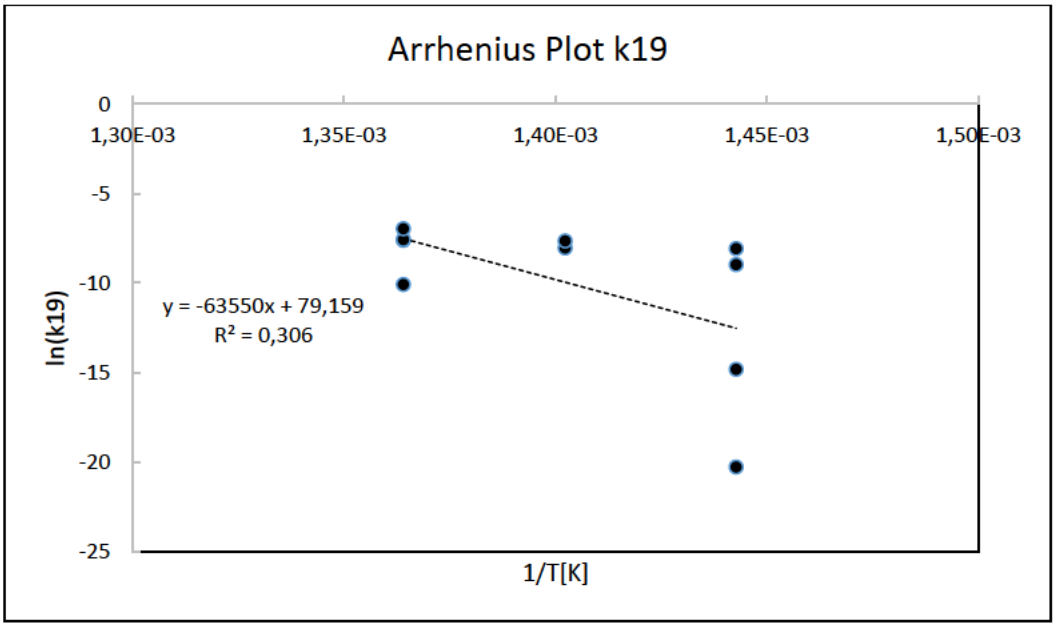


Figure 5-22. Arrhenius Plot for Parameter k19

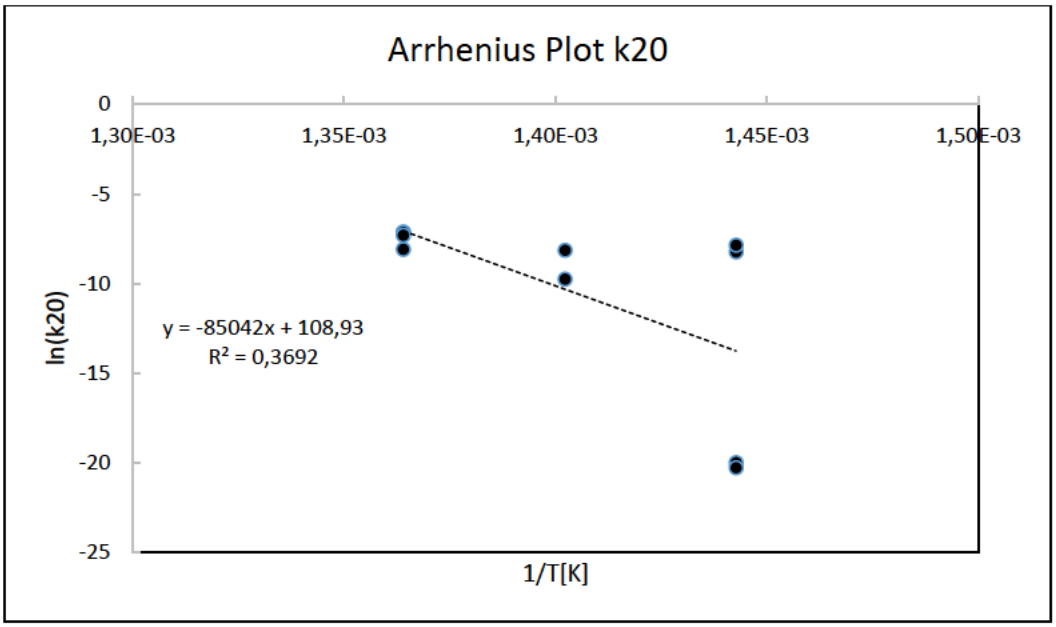


Figure 5-23. Arrhenius Plot for Parameter k20

Arrhenius plots – k21, k22, k23, k24 & k25

Significant spread exists in parameter *k21* when fitted for the experimental datasets. This is illustrated in Figure 5-24, which indicates a coefficient of determination (R^2) value of 0.4041 – This indicates that variance in the actual data is partially accounted for by the linear model, since it is above the selected threshold value. Therefore, the values shown in Table 5-3 would be more likely to be a suitable representation of reality, as the variance across all the investigated temperatures are adequately accounted for.

This implies that the constants in Table 5-3, determined for group E attaching to group A, would likely be an adequate representation of reality. Therefore, this will be used as a benchmark to verify the applicability of the size-based effect of the participating groups, where the existing group is E (parameter *k21* to *k25*). This will only be considered where R^2 values are above the selected relative threshold of 0.3. Considering that the values observed for parameter *k21* are applicable to the smallest (E) and largest (A) reacting groups, the assumption requires that all the determined constants for the applicable parameters be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k21*.

For the regression analysis performed in Figure 5-24 to 5-28, the detailed statistical parameters are provided in Table 5-3. Considering that this is based on the *Arrhenius transformation* of Section 2.13, the independent variable in the regression would be *1/T*, and the dependent variable would be *ln(k)*.

Similarly, R^2 values were 0.1769, 0.0034, 0.1199, and 0.0213 for parameters *k22* in Figure 5-25, *k23* in Figure 5-26, *k24* in Figure 5-27, and *k25* in Figure 5-28, respectively. In this parameter set, group E is the attaching group, while the existing group is varied, as illustrated in Table 5-3. From the selected relative threshold of 0.3, the linear models for parameters *k22*, *k23, k24,* and *k25*, do not adequately capture the variance in the actual data since it is below the selected relative threshold; therefore, the identified rate parameters in Table 5-3 will not be a good relative representation of reality.

The values observed for parameter *k22* are applicable to the smallest (E) and a moderately large (B) group, parameter *k23* are applicable to the smallest (E) and a medium sized (C) group, parameter *k24* are applicable to the smallest (E) and a moderately small (D) group, and parameter *k25* are applicable to the smallest groups (E & E), the resulting Arrhenius constants should be lower for activation energy and higher for the pre-exponential factor when compared to the values determined for parameter *k21*. However, these will not necessarily be observed due to the poor relative fit for these parameters.

Despite the potential inaccuracies, the activation energy value changes from a positive (endothermic) to a negative (exothermic) value, when *k23* is compared to *k21*. This indicates that there is a potential chemical effect that occurs when the groups defining *k23* interact and illustrates that there is likely more than steric factors that can hinder or promote the reaction.

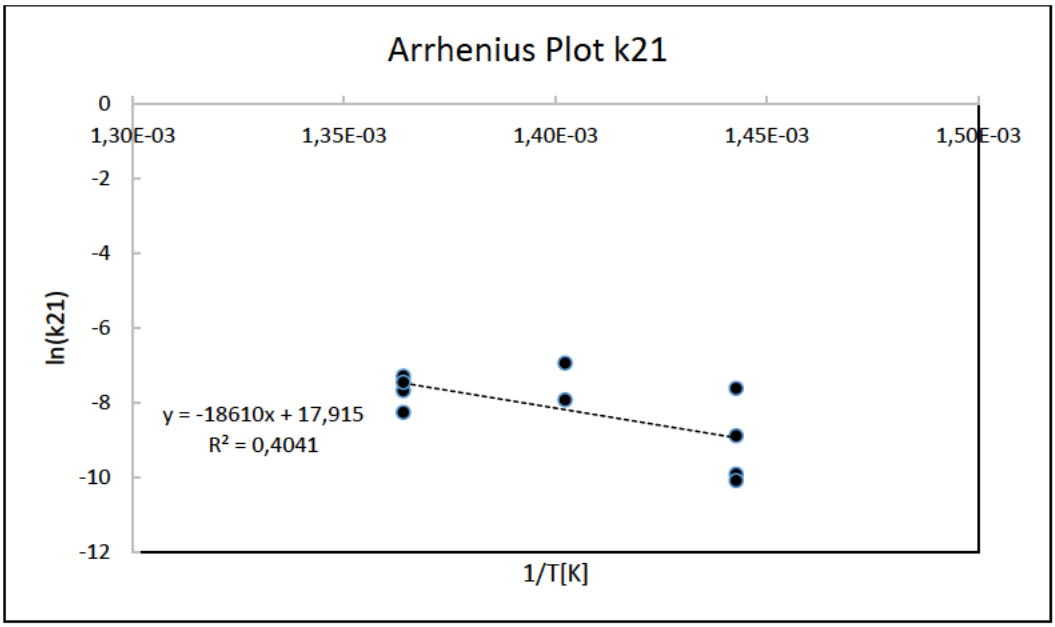


Figure 5-24. Arrhenius Plot for Parameter k21

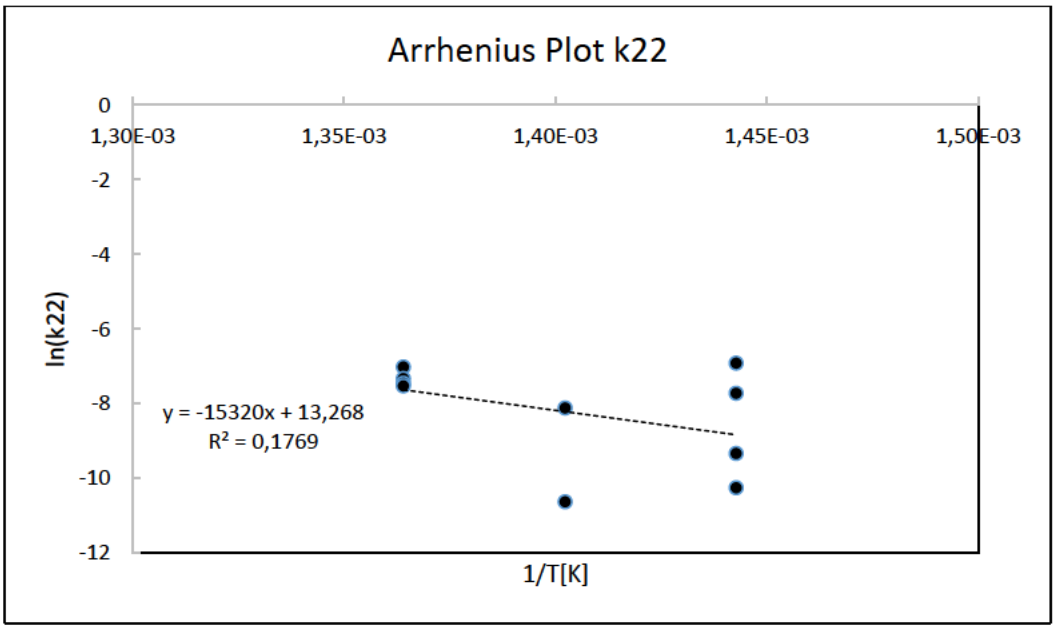


Figure 5-25. Arrhenius Plot for Parameter k22

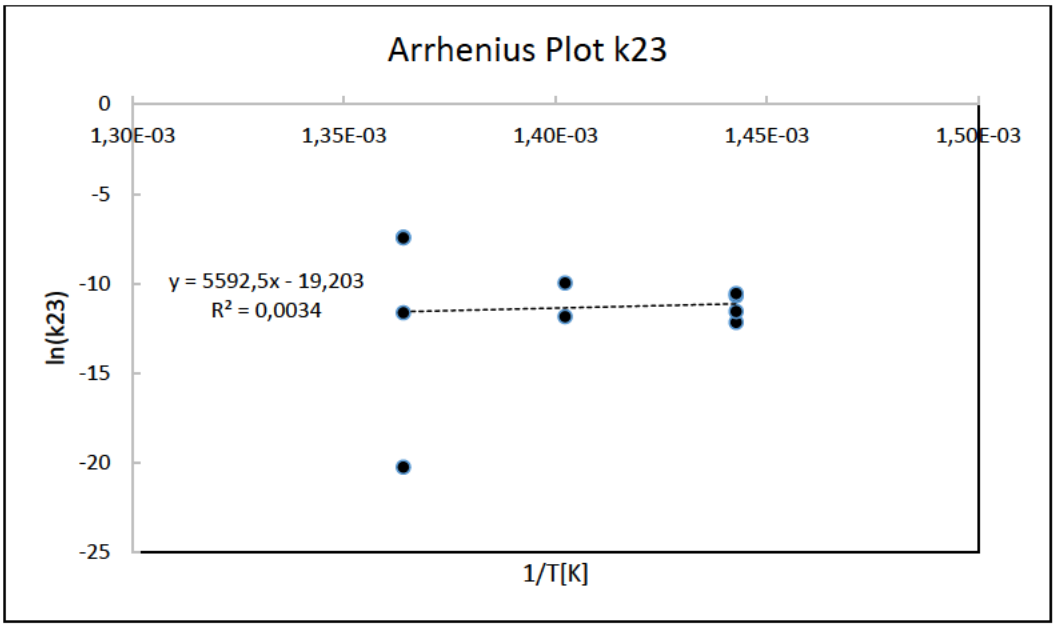


Figure 5-26. Arrhenius Plot for Parameter k23

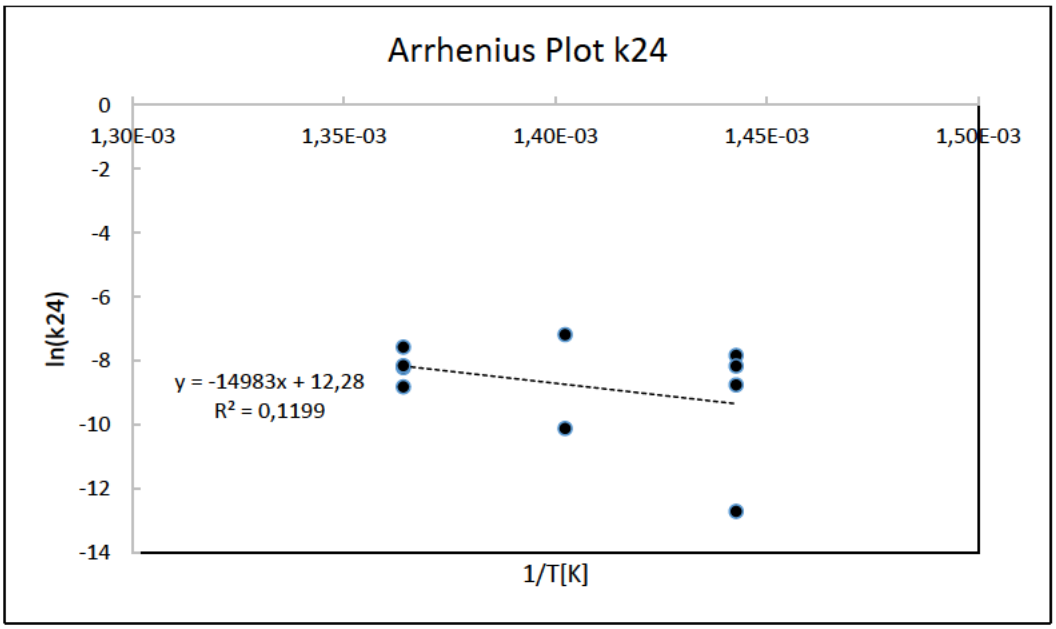


Figure 5-27. Arrhenius Plot for Parameter k24

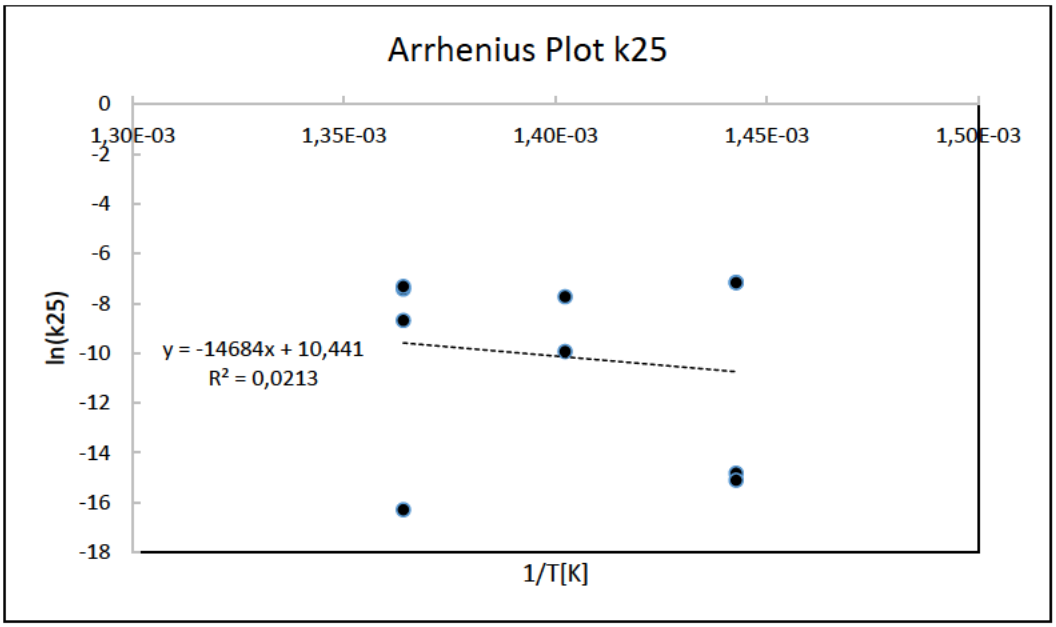


Figure 5-28. Arrhenius Plot for Parameter k25

5.7. Model Testing & Implementation

In the Arrhenius plot discussions, the results were compared on a relative basis. This was performed due to the similarities that were expected for cases involving the same attaching group. This allowed for normalising of the basis of comparison; however, individual parameter predictions still appeared to hold significant variability. Despite, the variability in statistical significance of the regression models fitted for the individual parameters, and the resulting constants, the optimisation approach was one that focused on *global minimisation* of the fitness function, of Figure 5-2.

Therefore, the set of all identified parameters should together capture the overall species interactions, within the temperature range investigated. Therefore, validation of this can be achieved by using the developed kinetic parameters, illustrated in Table 5-3, to predict the species flowrates, at the experimental conditions shown in Table A-1.

The initiating step in this process would be to substitute the individual parameter inputs, i.e., *Input 2* in the process of Figure D-1, with the kinetic model described by the activation energy and pre-exponential factor pairs shown in Table 5-3, using the following equation,

$$k_i = A_{o,i} e^{\left(\frac{-E_{a,i}}{RT}\right)}$$

Where k_i represents the *ith* rate parameter, and $A_{o,i}$ & $E_{a,i}$ are the respective pre-exponential factor and activation energy from Table 5-3. The code script adjusted to utilise the kinetic model is provided in Appendix E5, and the resulting predictions for each experimental dataset shown in Table A-2, at the conditions in Table A-1, are shown in Table 5-4.

Table 5-4. Kinetic Model Predictions & Fitness Values for Experimental Conditions in Table A-1

	Run 1 $\left[\frac{mol}{s}\right]$		Run 2 $\left[\frac{mol}{s}\right]$		Run 3 $\left[\frac{mol}{s}\right]$		Run 4 $\left[\frac{mol}{s}\right]$		Run 5 $\left[\frac{mol}{s}\right]$	
Product	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**
C_2	4,219E-07	4,219E-07	1,807E-06	1,807E-06	5,498E-07	5,498E-07	1,387E-06	1,387E-06	7,191E-07	7,191E-07
C_3	8,350E-07	8,350E-07	5,900E-06	5,900E-06	1,111E-06	1,111E-06	4,564E-06	4,564E-06	3,028E-06	3,028E-06
C_4	2,291E-07	2,291E-07	3,217E-06	3,217E-06	3,062E-07	3,062E-07	2,481E-06	2,481E-06	1,282E-06	1,282E-06
C_5	1,572E-06	1,572E-06	9,292E-06	9,292E-06	2,087E-06	2,087E-06	7,178E-06	7,178E-06	4,975E-06	4,975E-06
C_6	4,707E-05	4,707E-05	2,737E-05	2,737E-05	5,889E-05	5,889E-05	2,137E-05	2,137E-05	2,445E-05	2,445E-05
C_7	6,450E-07	6,450E-07	5,188E-06	5,188E-06	8,446E-07	8,446E-07	4,012E-06	4,012E-06	1,880E-06	1,880E-06
C_8	2,030E-07	2,030E-07	3,123E-06	3,123E-06	2,690E-07	2,690E-07	2,414E-06	2,414E-06	6,163E-07	6,163E-07
C_9	1,317E-07	1,317E-07	4,869E-06	4,869E-06	1,747E-07	1,747E-07	3,754E-06	3,754E-06	9,663E-07	9,663E-07
C_{10}	5,259E-07	5,259E-07	4,147E-06	4,147E-06	6,760E-07	6,760E-07	3,209E-06	3,209E-06	1,313E-06	1,313E-06
C_{11}	0	0	0	0	0	0	0	0	0	0
C_{12}	0	0	0	0	0	0	0	0	0	0
C_{13}	0	0	0	0	0	0	0	0	0	0
C_{14}	0	0	0	0	0	0	0	0	0	0
C_{15}	0	0	0	0	0	0	0	0	0	0
C_{16}	0	0	0	0	0	0	0	0	0	0
Fitness	4,389	5,370	1,968	2,640	37,40	50,55	2,231	3,095	4,853	9,121

The results in Table 5-4 were presented using parity plots, and these are shown in Figures 5-29 to 5-33 – In these figures, the kinetic model predicted flowrates, were plotted against the experimentally observed flowrates, shown in Table A-2, for all species. From Figures 5-29 to 5-33, the fit on all species flowrates is relatively good, as there is a very close scatter when compared to the $y=x$ line. This can also be observed in the fitness function values shown in Table 5-4 – The poorest prediction is observed in the prediction for *Run 3*, which shows the highest fitness function values, and the most scatter from the $y=x$ line.

Therefore, the overall kinetic model, and the developed modelling framework, adequately captures the interactions between the species, within the temperature range investigated. Even though the identified rate parameters exist in the lumped states shown in Appendix C5-2, the result of this work is clear – The modeling methodology proposed in *hypothesis 2*, and the viability of the group contribution approach has been confirmed.

As identified in Section 2.17, the existence of this approach represents an opportunity for extension into RCM kinetic modeling, where the application of a methodology similar to *hypothesis 2*, based on the interchange mechanism, could combine quantitative effectiveness, with theoretical soundness.

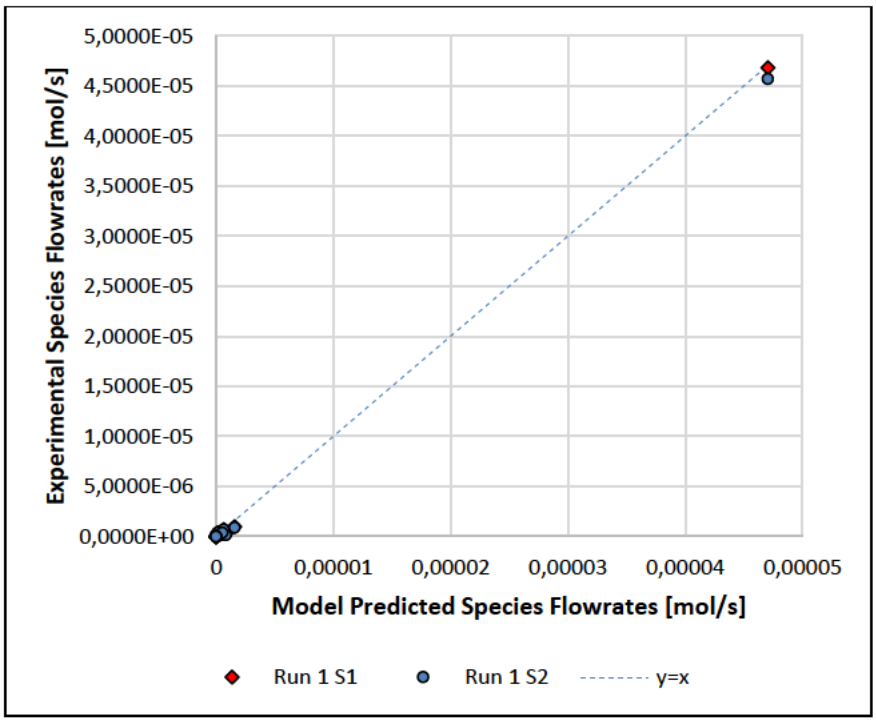


Figure 5-29. Observed (Experimental) vs Predicted (Model) Flowrate Parity Plot for Run 1 Data

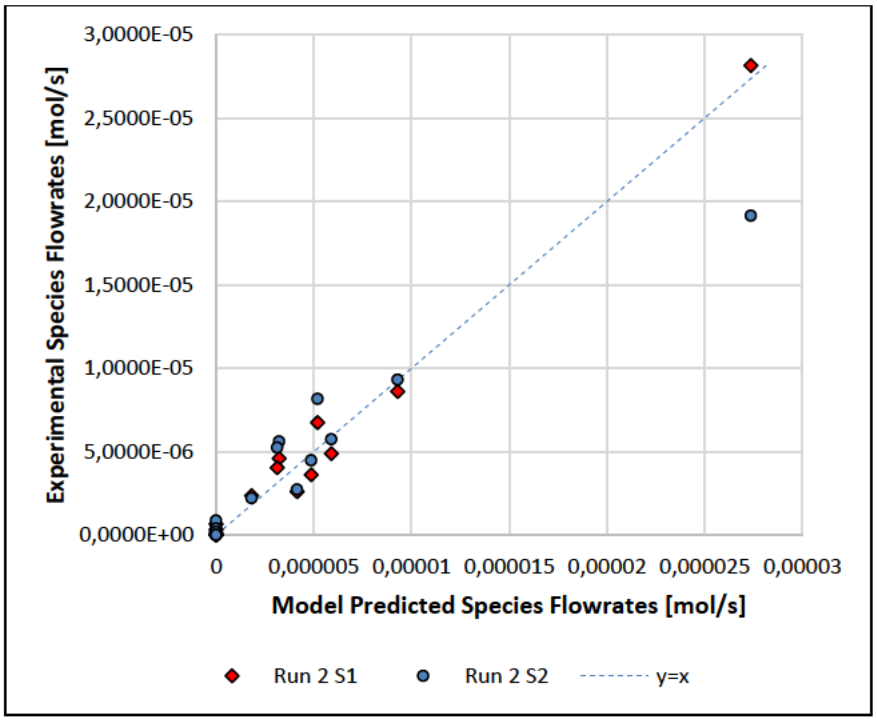


Figure 5-30. Observed (Experimental) vs Predicted (Model) Flowrate Parity Plot for Run 2 Data

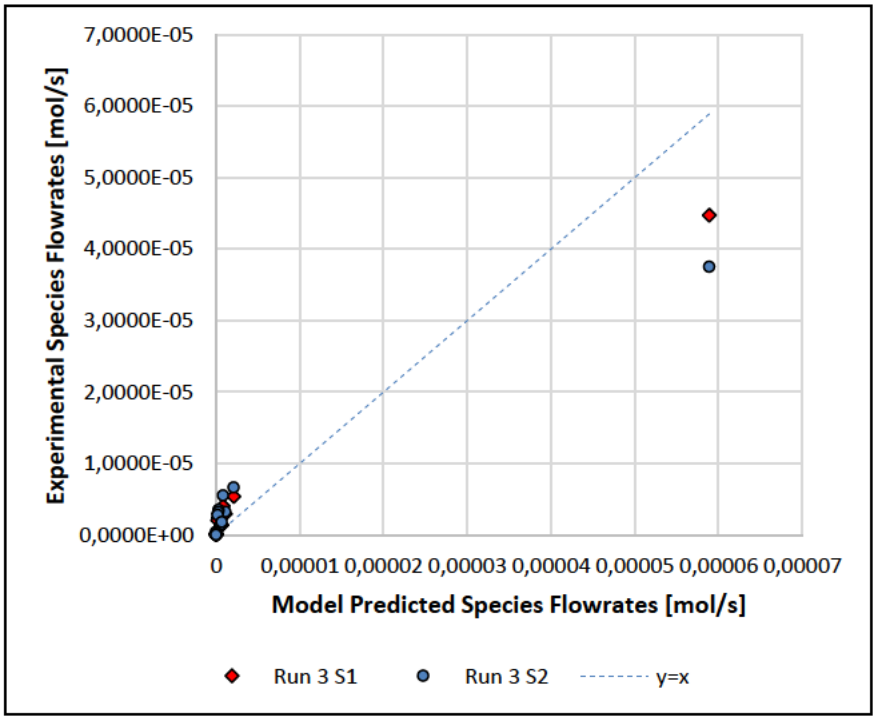


Figure 5-31. Observed (Experimental) vs Predicted (Model) Flowrate Parity Plot for Run 3 Data

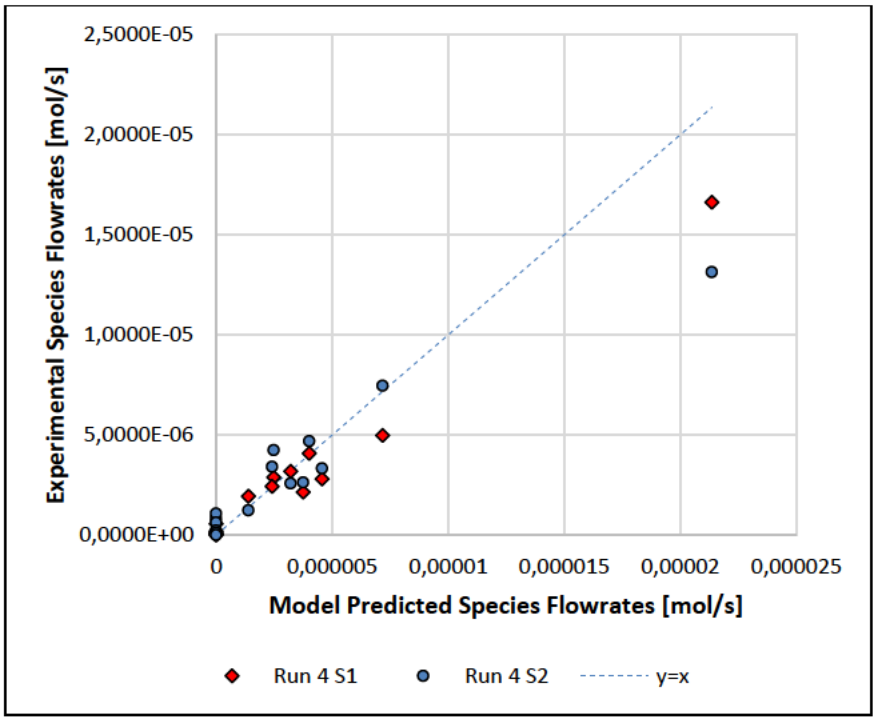


Figure 5-32. Observed (Experimental) vs Predicted (Model) Flowrate Parity Plot for Run 4 Data

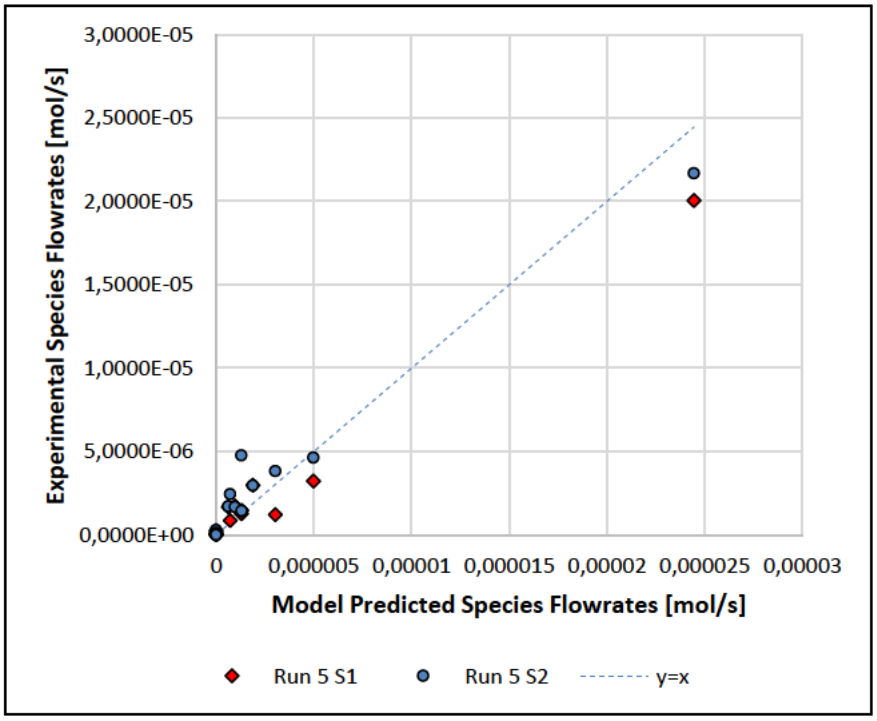


Figure 5-33. Observed (Experimental) vs Predicted (Model) Flowrate Parity Plot for Run 5 Data

6. Conclusion

The focus of this work was to develop and implement a novel group contribution kinetic modeling approach that could be generalized across the alkene metathesis landscape. Experimental data was obtained from literature, and attempts were made to isolate the metathesis effects by accounting for the isomerization of 1-Hexene, which was known to broaden the overall product distribution. Therefore, a key limitation in the applicability of the model becomes clear – In order to use the developed modeling methodology in practice, the chemistry of the system must be well understood to ensure that any competitive side reactions are accounted for. Side reactions are a very likely occurrence as temperatures increase and are associated with the variability of feedstock; however, focusing on more green synthesis methods will minimize the associated risk as these processes operate at much lower temperatures.

The focus in developing this modeling approach was on the generalized alkyl groups that appear within the metathesis mechanism, with the overall species used to understand driving forces and energetic effects – To capture these effects a framework was developed based on the linear metathesis mechanism, and key observations were made regarding the size of the groups involved. From literature, the size of groups and steric interactions between the groups were hypothesized to influence the reaction rate in different ways; however, the reaction needed to be accounted for at an elementary level, which was a key limitation in many of the developed hypotheses. After developing the reaction rate expressions, and implementing the hypotheses in simulation, it became clear (theoretically and numerically) that defining the rate parameters according to the existing group on the metallacyclobutane (transition state) complex, and the largest externally attaching alkyl group, yielded the most theoretically sound and stable system.

The quality of the determined kinetic parameters could have been improved if more than 10 carbons were accounted for in the generated reaction network; however, this was sufficient to understand the interactions occurring in the system, and for ascertaining the effectiveness of the novel group contribution kinetic modeling approach that was proposed. Overall, the modeling approach was found to be suitable as it resulted in predicted species flowrates that very closely compared to the experimentally observed species flowrates.

The methodology and approach developed in this work will allow for the design of systems that can convert petroleum process effluent streams (feedstock) into higher value (longer chain) chemical components, that can be used to create a synergistic industrial landscape by leveraging the benefits of the metathesis reaction. Future areas of work would involve applying the developed methodology to ring closing metathesis, as this has the specific ability to create industrial partnerships between the fine chemical (pharmaceutical) and petroleum industry.

7. References

- Alkorta, I. and Elguero, J., 2006. The carbon–carbon bond dissociation energy as a function of the chain length. *Chemical Physics Letters*, 425(4-6), pp.221-224.
- Bruner, R. (2020). The E-Z system for naming alkenes. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Organic_Chemistry/Supplemental_Modules_(Organic_Chemistry)/Fundamentals/Structure_of_Organic_Molecules/The_E-Z_system_for_naming_alkenes [Accessed 20 Feb. 2020].
- Cao, C. and Yuan, H., 2003. A New Approach of Evaluating Bond Dissociation Energy from Eigenvalue of Bonding Orbital-Connection Matrix for C−C and C−H Bonds in Alkane. *Journal of Chemical Information and Computer Sciences*, 43(2), pp.600-608.
- Cchange. (2020). *RSA Olefins | cChange*. [online] Available at: https://www.cchange.ac.za/rsa-olefins/ [Accessed 6 Jan. 2020].
- Clark, J. and Nguyen, T. (2020). *Physical Properties of Alkenes*. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Organic_Chemistry/Supplemental_Modules_(Organic_Chemistry)/Alkenes/Properties_of_Alkenes/Physical_Properties_of_Alkenes [Accessed 23 Jan. 2020].
- Cornaro, U., 1989. The catalytic effect of boron substitution in ZSM-5 type zeolites. Journal of Catalysis, 120(1), pp.182-191. doi.org/10.1016/0021-9517(89)90259-5.
- Dao, P., Latif, B. and Zhao, L., 2020. 9.7: Theories Of Reaction Rates. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Map%3A_Physical_Chemistry_for_the_Biosciences_(Chang)/09%3A_Chemical_Kinetics/9.07%3A_Theories_of_Reaction_Rates [Accessed 2 June 2020].
- Donde R. Anderson, Thay Ung, Garik Mkrtumyan, Guy Bertrand, Robert H. Grubbs, and Yann Schrodi. Organometallics 2008 27 (4), 563-566, DOI: 10.1021/om7008028
- Dos Santos, E., Granato, A. and Santos, A. (2018). Chapter 9. Metal-catalysed Metathesis Reactions for Greener Synthon/Drug Synthesis. *Green Chemistry Series*, pp.230-252.
- Fandrick, K., Savoie, J., Yee, N., Song, J. and Senanayake, C. (2014). Challenges and Opportunities for Scaling the Ring-Closing Metathesis Reaction in the Pharmaceutical Industry. *Olefin Metathesis*, pp.349-365.
- Fleming, P., 2020. 11.10: Collision Theory. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Book%3A_Physical_Chemistry_(Fleming)/11%3A_Chemical_Kinetics_I/11.10%3A_Collision_Theory [Accessed 2 June 2020].

- Fogler, H., 2006. Elements Of Chemical Reaction Engineering. Upper Saddle River, NJ: Pearson Education Internat, p.92.
- Froment, G. (2005). Single Event Kinetic Modeling of Complex Catalytic Processes. *Catalysis Reviews*, 47(1), pp.83-124.
- Grubbs, R. (2006). Olefin-Metathesis Catalysts for the Preparation of Molecules and Materials (Nobel Lecture). Angewandte Chemie International Edition, 45(23), pp.3760-3765.
- Hunt, I. (2020). E- and Z-alkenes. [online] Chem.ucalgary.ca. Available at: http://www.chem.ucalgary.ca/courses/351/WebContent/orgnom/alkenes/alkenes-03.html [Accessed 20 Feb. 2020].
- Hunter, K. and East, A., 2002. Properties of C−C Bonds inn-Alkanes: Relevance to Cracking Mechanisms. *The Journal of Physical Chemistry A*, 106(7), pp.1346-1356.
- Jacobsen, Ø., Klaveness, J. and Rongved, P. (2010) 'Structural and pharmacological effects of ring-closing metathesis in peptides', Molecules, 15(9), pp. 6638–6677. doi:10.3390/molecules15096638.
- Kapteijn, F. and Mol, J.C. (1982) 'Stereochemistry in metathesis of N-alkenes using heterogeneous oxide catalysts', Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases, 78(8), p. 2583. doi:10.1039/f19827802583.
- Katoch, S., Chauhan, S. S., & Kumar, V. (2020). A review on genetic algorithm: past, present, and future. Multimedia Tools and Applications. doi:10.1007/s11042-020-10139-6
- Laidler, K. and King, M., 1983. Development of Transition-state Theory. Chemischer Informationsdienst, 14(47), pp.2657-2664.
- La-Salvia, N., Lovón-Quintana, J., Lovón, A. and Valença, G., 2017. Influence of Aluminum Addition in the Framework of MCM-41 Mesoporous Molecular Sieve Synthesized by Non-Hydrothermal Method in an Alkali-Free System. Materials Research, 20(6), pp.1461-1469. doi.org/10.1590/1980-5373-mr-2016-1064.
- Levenspiel, O., 1999. Chemical Reaction Engineering. New York: Wiley, pp.13-33.
- LibreTexts. (2019). 8.7 Lewis Acids and Bases. [online] Available at: https://chem.libretexts.org/Bookshelves/General_Chemistry/Map%3A_Chemistry_(Averill_and_Eldredge)/08%3A_Ionic_versus_Covalent_Bonding/8.7_Lewis_Acids_and_Bases [Accessed 10 Jan. 2020].
- Lokhat, D., Starzak, M. and Stelmachowski, M., 2008. Gas-phase metathesis of 1-hexene over a WO3/SiO2 catalyst: Search for optimal reaction conditions. *Applied Catalysis A: General*, 351(2), pp.137-147.
- Louie, K., 2020. *Eyring Equation*. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Kinetics/Modeling_Reaction_Kinetics/Transition_State_Theory/Eyring_equation [Accessed 5 September 2020].
- Lundberg, D. (2006). The coordination chemistry of solvated metal ions in DMPU.

- Marshall, E. (2014). 313 Advanced Organometallics. [image] Available at: https://slideplayer.com/slide/2400850/ [Accessed 24 Feb. 2020].
- Mirjalili, S., Song Dong, J., Lewis, A., & Sadiq, A. S. (2019). Particle Swarm Optimization: Theory, Literature Review, and Application in Airfoil Design. Springer Series on Fluorescence, 167–184. doi:10.1007/978-3-030-12127-3_10
- Moravec, J. (2015). A Comparative Study: L1-Norm Vs. L2-Norm; Point-to-Point Vs. Point-to-Line Metric; Evolutionary Computation Vs. Gradient Search. Applied Artificial Intelligence. 29. 10.1080/08839514.2015.993560.
- Nelson, D.J. *et al.* (2011) 'Toward a simulation approach for alkene ring-closing metathesis: Scope and limitations of a model for RCM', The Journal of Organic Chemistry, 76(20), pp. 8386–8393. doi:10.1021/jo201611z.
- Nguyen, T. and Ngo, M., 2020. Elementary Reactions. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Kinetics/Rate_Laws/Reaction_Mechanisms/Elementary_Reactions [Accessed 2 June 2020].
- Perez-Benito, J., 2017. Some Considerations on the Fundamentals of Chemical Kinetics: Steady State, Quasi-Equilibrium, and Transition State Theory. *Journal of Chemical Education*, 94(9), pp.1238-1246.
- Petrik, L. (2009). The influence of cation, anion and water content on the rate of formation and pore size distribution of zeolite ZSM-5. South African Journal of Science. 105. 251-257. doi.org/10.4102/sajs.v105i7/8.37.
- Prithipal, A., Lokhat, D. and Starzak, M. (2015). Metathesis of 1-Hexene over Heterogeneous Tungsten-Based Catalysts. *Chemical Engineering & Technology*, 38(8), pp.1343-1352.
- Ptáček, P., Šoukal, F. and Opravil, T. (2020). Introduction to the Transition State Theory. [online] Available at: http://dx.doi.org/10.5772/intechopen.78705 [Accessed 20 Feb. 2020].
- Rodríguez-González, L., Hermes, F., Bertmer, M., Rodríguez-Castellón, E., Jiménez-López, A. and Simon, U., 2007. The acid properties of H-ZSM-5 as studied by NH3-TPD and 27Al-MAS-NMR spectroscopy. Applied Catalysis A: General, 328(2), pp.174-182. doi.org/10.1016/j.apcata.2007.06.003.
- Schaller, C. (2020). *11.6: Alkene Metathesis*. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Organic_Chemistry/Map%3A_Organic_Chemistry_(Bruice)/11%3A_Organometallic_Compounds/11.06%3A_Alkene_Metathesis [Accessed 19 Feb. 2020].
- Schrock, R. and Hoveyda, A. (2003). Molybdenum and Tungsten Imido Alkylidene Complexes as Efficient Olefin-Metathesis Catalysts. Angewandte Chemie International Edition, 42(38), pp.4592-4633.
- Sevy, A., Huffaker, R. and Morse, M., 2017. Bond Dissociation Energies of Tungsten Molecules: WC, WSi, WS, WSe, and WCl. *The Journal of Physical Chemistry A*, 121(49), pp.9446-9457.

- Simoes, J. and Beauchamp, J., 1990. Transition metal-hydrogen and metal-carbon bond strengths: the keys to catalysis. *Chemical Reviews*, [online] 90(4). Available at: https://pubs.acs.org/doi/abs/10.1021/cr00102a004 [Accessed 26 November 2020].
- Sinclair, F. *et al.* (2017) 'Olefin cross metathesis and ring-closing metathesis in polymer chemistry', Polymer Chemistry, 8(22), pp. 3385–3398. doi:10.1039/c7py00340d.
- Skinner, H. and Connor, J., 1985. Metal-ligand bond-energies in organometallic compounds. *Pure and Applied Chemistry*, 57(1), pp.79-88.
- Smith, J., Van Ness, H. and Abbott, M., 2001. *Introduction To Chemical Engineering Thermodynamics*. 6th ed. New York: McGraw-Hill, pp.450-464.
- Speight, J., Lange, N. and Dean, J., 2005. *Lange's Handbook Of Chemistry*. 16th ed. New York, USA: McGraw-Hill Professional Publishing, pp.4.43.
- Sundaramurthy, V., 2000. Synthesis Characterisation And Catalytic Properties Of B ZSM 5 And B MCM 41 Molecular Sieves. Ph.D. Anna University. Viewed 18 March 2020. http://hdl.handle.net/10603/75127
- Svoboda, G., Vynckier, E., Debrabandere, B. and Froment, G. (1995). Single-Event Rate Parameters for Paraffin Hydrocracking on a Pt/US-Y Zeolite. *Industrial & Engineering Chemistry Research*, 34(11), pp.3793-3800.
- Teske, S. *et al.* (2022) 'Renewable Energy for Industry Supply', Achieving the Paris Climate Agreement Goals, pp. 225–246. doi:10.1007/978-3-030-99177-7_9.
- Toreki, R. (2002). The Organometallic HyperTextBook: Olefin Metathesis. [online] Ilpi.com. Available at: http://www.ilpi.com/organomet/olmetathesis.html [Accessed 21 Feb. 2020].
- Trong On, D., Joshi, P., Lemay, G. and Kaliaguine, S., 1995. Acidity and structural state of boron in mesoporous boron silicate MCM-41. Zeolites: A Refined Tool for Designing Catalytic Sites, Proceedings of the International Zeolite Symposium, pp.543-549. doi.org/10.1016/s0167-2991(06)81937-3.
- Truhlar, D. and Kohen, A., 2001. Convex Arrhenius plots and their interpretation. Proceedings of the National Academy of Sciences, 98(3), pp.848-851.
- Tu, C., Li, M., Li, H., Chu, Y., Liu, F., Nie, H. and Li, D., 2016. Effects of sulfur compounds on the hydrogenation and isomerization of 1-hexene over a sulfided CoMo catalyst for hydrodesulfurization. RSC Advances, 6(39), pp.33177-33183.
- Żukowska, K., Szadkowska, A. and Grela, K. (2013) 'Olefin metathesis', Comprehensive Inorganic Chemistry II, pp. 105–126. doi:10.1016/b978-0-08-097774-4.00606-9.

8. Acknowledgements

Thank you to my supervisor, Professor David Lokhat, and my family & friends for their support, guidance, and motivation over the last few years.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged and appreciated.

Appendix A: Literature Experimental Data

Table A-1. Experimental Data [1/2] (Prithipal *et al.*, 2015)

	Run 1	**Run 2**	**Run 3**	**Run 4**	**Run 5**
Pressure [kPa]	98,8	98	99,1	98,4	98,5
Temperature [°C]	420	460	420	460	440
Feed Composition [mol % 1-hexene]	59,99	78,00	78,80	59,81	71,92
Space Time [g.min.mol^{-1}]	193,48	200,26	404,60	394,15	305,53
X [%]	10,39	63,57	36,69	70,44	46,84
Yield C_{10}-C_{16} [%]	0,72	6,13	3,28	9,57	4,85
Yield C_{10} [%]	0,70	4,10	2,47	5,72	3,71
W_{cat} [g]	1	1	2	2	1
Highest Length Carbon [in products]	10	14	14	14	14
$F_{1\text{-Hex in}}$ [g/min]	0,2607	0,3278	0,3278	0,2543	0,1981
$F_{1\text{-Hex in}}$ [mol/s]	5,163E-05	6,491E-05	6,491E-05	5,036E-05	3,923E-05
F_{N2} [mol/s]	3,444E-05	1,831E-05	1,746E-05	3,384E-05	1,532E-05
$F_{tot\ in}$ [mol/s]	8,607E-05	8,322E-05	8,237E-05	8,421E-05	5,455E-05
$Q_{tot\ in}$ [m^3/s]	5,020E-06	5,176E-06	4,790E-06	5,216E-06	3,284E-06
$C_{1\text{-Hex tot in}}$ [mol/m^3]	10,285	12,541	13,551	9,655	11,948

Table A-2. Experimental Data [2/2] (Prithipal *et al.*, 2015)

	Run 1 $\left[\frac{mol}{s}\right]$		**Run 2** $\left[\frac{mol}{s}\right]$		**Run 3** $\left[\frac{mol}{s}\right]$		**Run 4** $\left[\frac{mol}{s}\right]$		**Run 5** $\left[\frac{mol}{s}\right]$	
Product	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**
C_2	2,631E-07	1,346E-07	2,375E-06	2,207E-06	1,349E-06	1,431E-06	1,948E-06	1,240E-06	8,606E-07	2,442E-06
C_3	3,163E-07	1,555E-07	4,876E-06	5,747E-06	2,937E-06	3,237E-06	2,813E-06	3,331E-06	1,221E-06	3,831E-06
C_4	1,702E-07	8,512E-08	4,601E-06	5,616E-06	2,959E-06	3,488E-06	2,881E-06	4,243E-06	1,277E-06	4,772E-06
C_5	9,853E-07	8,696E-07	8,600E-06	9,318E-06	5,385E-06	6,651E-06	4,983E-06	7,467E-06	3,232E-06	4,622E-06
C_6	4,682E-05	4,572E-05	2,815E-05	1,915E-05	4,470E-05	3,750E-05	1,663E-05	1,314E-05	2,005E-05	2,167E-05
C_7	6,851E-07	6,805E-07	6,747E-06	8,159E-06	4,036E-06	5,522E-06	4,080E-06	4,689E-06	2,972E-06	2,952E-06
C_8	3,214E-07	3,244E-07	4,034E-06	5,239E-06	2,363E-06	3,202E-06	2,431E-06	3,411E-06	1,678E-06	1,721E-06
C_9	3,640E-07	3,782E-07	3,610E-06	4,483E-06	2,021E-06	2,802E-06	2,145E-06	2,630E-06	1,725E-06	1,674E-06
C_{10}	3,429E-07	3,815E-07	2,596E-06	2,726E-06	1,396E-06	1,807E-06	3,186E-06	2,575E-06	1,460E-06	1,451E-06
C_{11}	0,000E+00	1,075E-08	6,649E-07	8,787E-07	2,454E-07	4,053E-07	5,625E-07	8,323E-07	2,560E-07	2,848E-07
C_{12}	0,000E+00	3,570E-09	3,300E-07	1,354E-08	1,793E-08	6,630E-09	2,046E-07	1,077E-06	2,186E-09	2,779E-09
C_{13}	0,000E+00	0,000E+00	9,892E-09	3,944E-07	1,015E-07	1,480E-07	1,152E-07	6,218E-07	1,113E-07	1,202E-07
C_{14}	0,000E+00	0,000E+00	1,480E-07	1,957E-07	5,320E-08	7,363E-08	5,631E-08	2,315E-07	5,603E-08	5,774E-08
C_{15}	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	5,252E-08	1,220E-07	0,000E+00	0,000E+00
C_{16}	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00

Appendix B1: 1-Hexene Isomerisation Network

	Reactions	Rate Expressions
1	1-Hexene (A) ⟷ 2-Hexene (B)	$r_{1f} = k_{1f}C_A$ and $r_{1r} = k_{1r}C_B$ Note: $k_{1r} = \frac{k_{1f}}{K_{1e}}$ and $C_j = \frac{F_j[mol/s]}{q_o[m^3/s]}$
2	1-Hexene (A) ⟷ 3-Hexene (C)	$r_{2f} = k_{2f}C_A$ and $r_{2r} = k_{2r}C_C$ Note: $k_{2r} = \frac{k_{2f}}{K_{2e}}$ and $C_j = \frac{F_j[mol/s]}{q_o[m^3/s]}$
3	2-Hexene (B) ⟷ 3-Hexene (C)	$r_{3f} = k_{3f}C_B$ and $r_{3r} = k_{3r}C_C$ Note: $k_{3r} = \frac{k_{3f}}{K_{3e}}$ and $C_j = \frac{F_j[mol/s]}{q_o[m^3/s]}$

Model Equations

$$\frac{dF_A}{d_W} = -(r_{1f} + r_{2f}) + (r_{1r} + r_{2r})$$

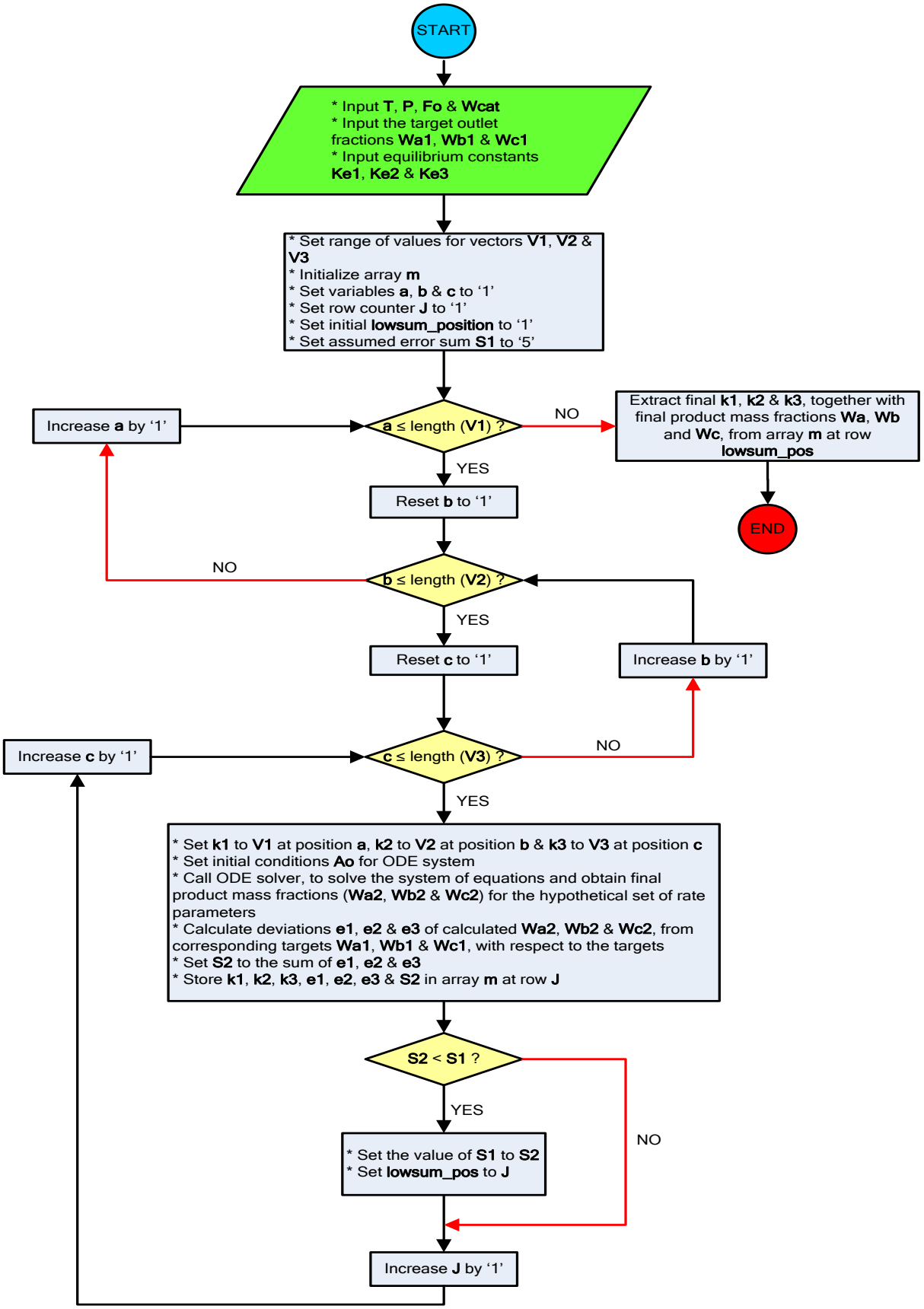
$$\frac{dF_B}{d_W} = -(r_{1r} + r_{3f}) + (r_{1f} + r_{3r})$$

$$\frac{dF_C}{d_W} = -(r_{2r} + r_{3r}) + (r_{2f} + r_{3f})$$

Notation

r_{if} – Forward reaction rate [mol/kg/s] of *ith* reaction
r_{ir} – Reverse reaction rate [mol/kg/s] of *ith* reaction
k_{if} – Forward reaction rate constant [m^3/kg/s] of *ith* reaction
k_{ir} – Reverse reaction rate constant [m^3/kg/s] of *ith* reaction
w – Mass of catalyst [kg]
F_j – *jth* component molar flowrate [mol/s]
C_j – *jth* component concentration [mol/m^3]
q_o – Bulk gas volumetric flowrate [m^3/s]

Appendix B2: 1-Hexene Isomerisation Rate Parameter Optimization Algorithm



Appendix B3: Predicted Optimal Reactor Profiles for 1-Hexene Isomerisation

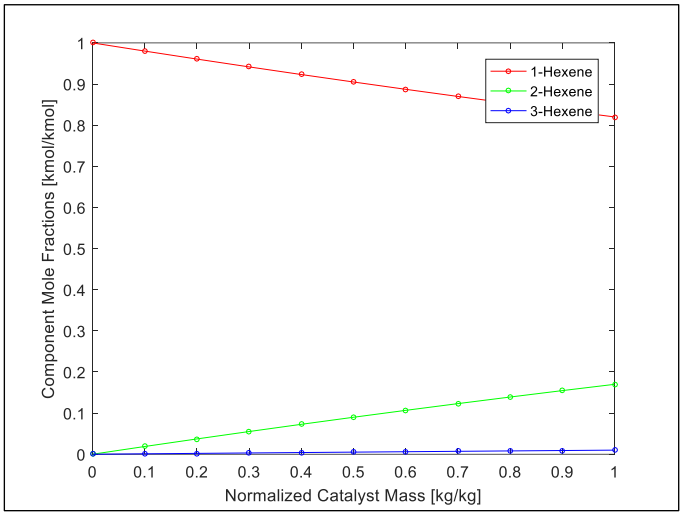


Figure B-1. Optimal Profile for T=473K and M_{cat}=5.2145g

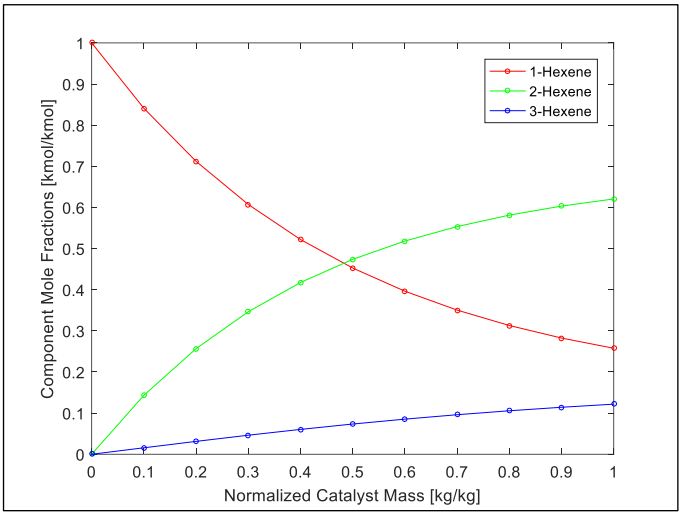


Figure B-3. Optimal Profile for T=573K and M_{cat}=5.2145g

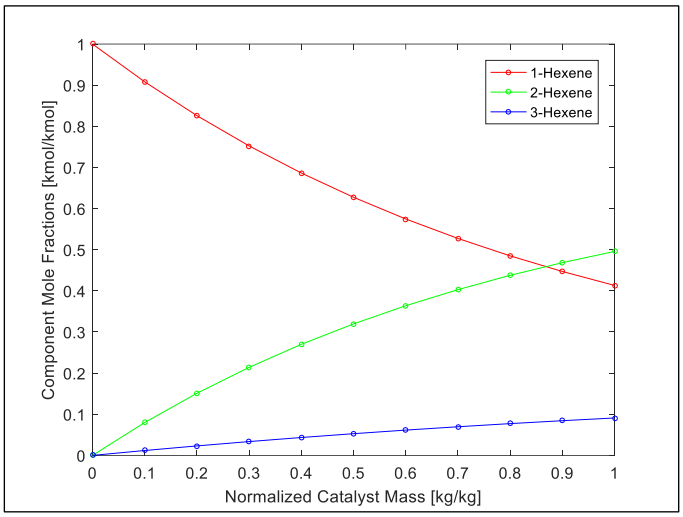


Figure B-2. Optimal Profile for T=523K and M_{cat}=5.2145g

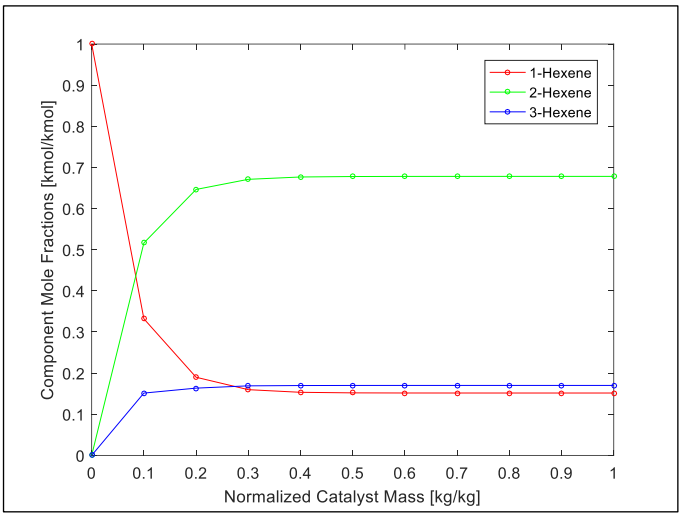
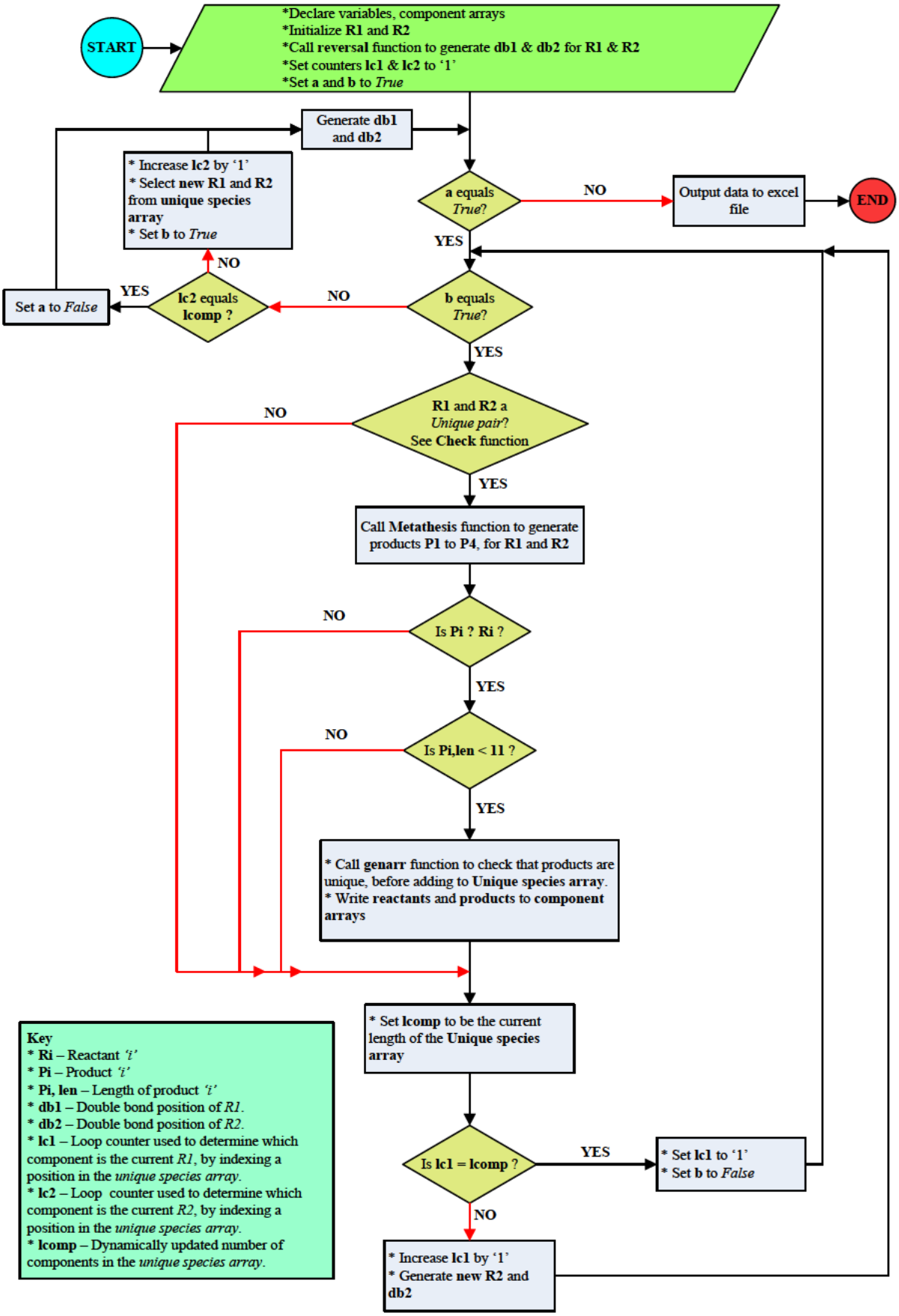


Figure B-4. Optimal Profile for T=623K and M_{cat}=5.2145g

Appendix C1: Network Generation Algorithm



Appendix C2: Reaction Network and Equilibrium Constants

Table C-1. Generated Metathesis Reaction Network and Equilibrium Constants

	Reactant 1	*Reactant 2*		*Product 1*	*Product 2*	*Equilibrium Constants*		
						693,15K	**713,15K**	**733,15K**
1	1-hexene	1-hexene	↔	1-ethene	5-decene	0,09038	0,09285	0,09529
2	1-hexene	2-hexene	↔	1-pentene	2-heptene	0,70919	0,70455	0,70022
3	1-hexene	2-hexene	↔	1-propene	4-nonene	0,93854	0,94382	0,94892
4	1-hexene	3-hexene	↔	1-butene	3-octene	0,76394	0,76243	0,76104
5	1-hexene	1-pentene	↔	1-ethene	4-nonene	0,08903	0,09138	0,09368
6	1-hexene	2-heptene	↔	1-propene	5-decene	1,34340	1,36125	1,37846
7	1-hexene	1-propene	↔	1-ethene	2-heptene	0,06727	0,06821	0,06913
8	1-hexene	4-nonene	↔	1-pentene	5-decene	1,01512	1,01616	1,01719
9	1-hexene	1-butene	↔	1-ethene	3-octene	0,10188	0,10455	0,10717
10	1-hexene	3-octene	↔	1-butene	5-decene	0,88706	0,88811	0,88915
11	2-hexene	3-octene	↔	2-heptene	3-heptene	0,35786	0,35016	0,34303
12	2-hexene	3-octene	↔	2-pentene	4-nonene	0,65507	0,64923	0,64376
13	2-hexene	2-hexene	↔	2-butene	4-octene	0,99313	0,99298	0,99285
14	2-hexene	3-hexene	↔	2-pentene	3-heptene	0,28897	0,28148	0,27457
15	2-hexene	1-ethene	↔	1-propene	1-pentene	10,54177	10,32892	10,12899
16	2-hexene	5-decene	↔	2-heptene	4-nonene	0,69863	0,69335	0,68839
17	2-hexene	1-pentene	↔	1-propene	4-octene	0,92457	0,92882	0,93288
18	2-hexene	2-heptene	↔	2-butene	4-nonene	1,42153	1,43215	1,44227
19	2-hexene	1-propene	↔	2-butene	1-pentene	1,07415	1,06909	1,06428
20	2-hexene	4-nonene	↔	2-heptene	4-octene	0,69863	0,69335	0,68839
21	2-hexene	1-butene	↔	2-pentene	1-pentene	0,74964	0,74285	0,73647
22	2-hexene	1-butene	↔	1-propene	3-heptene	0,54196	0,53671	0,53180
23	2-hexene	3-heptene	↔	2-pentene	4-octene	1,27888	1,28556	1,29190
24	2-hexene	2-pentene	↔	2-butene	3-heptene	0,77656	0,77241	0,76852
25	3-hexene	4-octene	↔	3-heptene	3-heptene	0,22596	0,21895	0,21253
26	3-hexene	1-ethene	↔	1-butene	1-butene	7,49814	7,29232	7,10097
27	3-hexene	5-decene	↔	3-octene	3-octene	0,86120	0,85848	0,85592
28	3-hexene	1-pentene	↔	3-heptene	1-butene	0,38548	0,37892	0,37282
29	3-hexene	2-heptene	↔	3-octene	2-pentene	0,80751	0,80386	0,80043
30	3-hexene	1-propene	↔	2-pentene	1-butene	0,53320	0,52446	0,51630
31	3-hexene	4-nonene	↔	3-octene	3-heptene	0,44113	0,43355	0,42651
32	3-hexene	2-butene	↔	2-pentene	2-pentene	0,37212	0,36441	0,35728
33	1-ethene	4-octene	↔	1-pentene	1-pentene	11,40177	11,12052	10,85775
34	1-ethene	3-heptene	↔	1-pentene	1-butene	19,45132	19,24503	19,04652

35	1-ethene	2-pentene	↔	1-butene	1-propene	14,06240	13,90454	13,75347
36	1-ethene	2-butene	↔	1-propene	1-propene	9,81406	9,66145	9,51722
37	5-decene	4-octene	↔	4-nonene	4-nonene	1,00000	0,99999	0,99999
38	5-decene	3-heptene	↔	4-nonene	3-octene	1,95226	1,98010	2,00679
39	5-decene	2-pentene	↔	3-octene	2-heptene	1,06649	1,06795	1,06932
40	5-decene	2-butene	↔	2-heptene	2-heptene	0,49146	0,48413	0,47730
41	1-pentene	2-heptene	↔	1-propene	4-nonene	1,32340	1,33960	1,35516
42	1-pentene	4-nonene	↔	1-hexene	4-octene	0,98511	0,98410	0,98310
43	1-pentene	3-octene	↔	1-hexene	3-heptene	0,50460	0,49699	0,48988
44	1-pentene	3-octene	↔	1-butene	4-nonene	0,87385	0,87398	0,87412
45	1-pentene	3-heptene	↔	1-butene	4-octene	1,70599	1,73059	1,75419
46	2-heptene	1-propene	↔	2-butene	1-hexene	1,51461	1,51740	1,51991
47	2-heptene	1-butene	↔	2-pentene	1-hexene	1,05704	1,05435	1,05176
48	2-heptene	1-butene	↔	1-propene	3-octene	1,51445	1,53276	1,55031
49	2-heptene	3-heptene	↔	2-pentene	4-nonene	1,83054	1,85412	1,87669
50	2-heptene	2-pentene	↔	2-butene	3-octene	2,17003	2,20591	2,24038
51	1-propene	3-octene	↔	1-hexene	2-pentene	0,69797	0,68788	0,67842
52	1-propene	3-heptene	↔	1-pentene	2-pentene	1,38321	1,38408	1,38485
53	1-propene	2-pentene	↔	1-butene	2-butene	1,43288	1,43918	1,44511
54	4-nonene	1-butene	↔	3-heptene	1-hexene	0,57744	0,56865	0,56043
55	4-nonene	3-heptene	↔	4-octene	3-octene	1,95227	1,98011	2,00680

Note: The reactions written in red text involve those reactants that can undergo metathesis to form two sets of product pairs.

Appendix C3: Equilibrium Constant Estimation Method

In all the equations in this section the *change* (Δ) refers to the following difference,

$$\Delta Z = \sum_i \nu_i Z_i - \sum_j \nu_j Z_j$$

Where, subscripts *i* and *j* refer to the *products* and *reactants* respectively, ν refers to the *stoichiometric coefficient* of a species and Z refers to some *thermodynamic property*.

According to *Smith, Van Ness and Abbott (2001),* the equilibrium constant (K_{eq}) is given by the following relation,

$$K_{eq} = e^{-\frac{\Delta G^o}{RT}}$$

Where, *T* is the temperature at which the reaction is occurring, ΔG^o is the *standard Gibbs-energy change* of the reaction at temperature *T* and *R* is the *universal gas constant*. Note that the units of the exponent must be dimensionless; therefore, the units of ΔG^o must be the same as the product *RT*.

$$\frac{\Delta G^o}{RT} = \frac{\Delta G^o_{fO} - \Delta H^o_{fO}}{RT_O} + \frac{\Delta H^o_{fO}}{RT} + \left(\frac{1}{T}\right)\left(\int_{T_O}^{T} \frac{\Delta C^o_p}{R}\, dT\right) - \left(\int_{T_O}^{T} \frac{\Delta C^o_p}{RT}\, dT\right)$$

Where, ΔG^o_{fO} is the *standard Gibbs-energy change of formation* at the reference state, ΔH^o_{fO} is the *standard Enthalpy change of formation* at the reference state and ΔC^o_p is the *standard change in Heat capacity*. *Smith, Van Ness and Abbott (2001)* provides a complete derivation of the expression above, which can be used to determine the value of the *standard Gibbs-energy change* (ΔG^o) of the reaction, at some temperature *T*. Note that the subscript *'o'* refers to the *gaseous reference state*, i.e., ideal gas at 1bar and 298.15K.

Smith, Van Ness and Abbott (2001) provided an expression for $\frac{\Delta C^o_p}{R}$ in the following form,

$$\frac{\Delta C^o_p}{R} = \Delta A + \Delta B \times T + \Delta C \times T^2 + \frac{\Delta D}{T^2}$$

Therefore, the integrals can be integrated from first principles as follows,

$$IDCPH = \int_{T_O}^{T} \frac{\Delta C^o_p}{R}\, dT = \Delta A(T - T_O) + \left(\frac{\Delta B}{2}\right)(T^2 - T_O^2) + \left(\frac{\Delta C}{3}\right)(T^3 - T_O^3) - \Delta D\left(\frac{1}{T} - \frac{1}{T_O}\right)$$

$$IDCPS = \int_{T_O}^{T} \frac{\Delta C^o_p}{RT}\, dT = \Delta A\, ln\left(\frac{T}{T_O}\right) + \Delta B(T - T_O) + \left(\frac{\Delta C}{2}\right)(T^2 - T_O^2) - \left(\frac{\Delta D}{2}\right)\left(\frac{1}{T^2} - \frac{1}{T_O^2}\right)$$

Using these integrals, the value of $\frac{\Delta G^o}{RT}$ can be determined; thus K_{eq} can be calculated at some temperature *T*. The data in Table C-2 was used to estimate the equilibrium constants of the reactions occurring in the metathesis network. Some of the required properties were not readily available in literature; therefore, the *Joback-Reid group contribution method* was used to estimate these properties.

The *A, B, C & D parameters* in Table C-2 represent the refitted parameters for the form provided in Smith, Van Ness and Abbott (2001). These parameters were generated using the *curve_fit* method, available in the *optimize* class of the *scipy* python library. The transformations are represented as follows,

$$\underbrace{\frac{C_p^o}{R} = W + XT + YT^2 + ZT^3}_{\boldsymbol{Joback - Reid\ Form}} \quad \rightarrow \quad \underbrace{\frac{C_p^o}{R} = A + BT + CT^2 + \frac{D}{T^2}}_{\boldsymbol{SVA\ Form}}$$

Therefore, both equations explain the same $\frac{C_p^o}{R}$ behavior with different algebraic forms and regression coefficients.

Table C-2. Pure Property Data for Metathesis Network Equilibrium Constant Predictions

Species	$\boldsymbol{C_p^o/R}$				$\boldsymbol{\Delta G_{fO}^o}$ **[J/mol]**	$\boldsymbol{\Delta H_{fO}^o}$ **[J/mol]**
	A	**B**	**C**	**D**		
1-Ethene	1,4240E+00	1,4394E-02	-4,3920E-06	0,0000E+00	6,8460E+04	5,2510E+04
1-Propene	1,6370E+00	2,2706E-02	-6,9150E-06	0,0000E+00	6,2205E+04	1,9710E+04
1-Butene	1,9670E+00	3,1630E-02	-9,8730E-06	0,0000E+00	7,0340E+04	-5,4000E+02
2-Butene	5,2350E+00	2,5583E-02	-7,0766E-06	-2,1209E+05	6,3020E+04	-1,0800E+04
1-Pentene	2,6910E+00	3,9753E-02	-1,2447E-05	0,0000E+00	7,8410E+04	-2,1280E+04
2-Pentene	6,1241E+00	3,3745E-02	-9,8832E-06	-2,4221E+05	7,1440E+04	-3,2200E+04
1-Hexene	3,2200E+00	4,8189E-02	-1,5157E-05	0,0000E+00	8,6830E+04	-4,1950E+04
2-Hexene	7,0131E+00	4,1908E-02	-1,2690E-05	-2,7232E+05	7,9860E+04	-5,1000E+04
3-Hexene	7,0131E+00	4,1908E-02	-1,2690E-05	-2,7232E+05	7,9860E+04	-4,9300E+04
2-Heptene	7,9022E+00	5,0070E-02	-1,5496E-05	-3,0244E+05	8,8280E+04	-7,3200E+04
3-Heptene	7,9022E+00	5,0070E-02	-1,5496E-05	-3,0244E+05	8,8280E+04	-7,3500E+04
3-Octene	8,7913E+00	5,8233E-02	-1,8303E-05	-3,3256E+05	9,6700E+04	-9,1230E+04
4-Octene	8,7913E+00	5,8233E-02	-1,8303E-05	-3,3256E+05	9,6700E+04	-9,1230E+04
4-Nonene	9,6803E+00	6,6395E-02	-2,1110E-05	-3,6268E+05	1,0512E+05	-1,1187E+05
5-Decene	1,0569E+01	7,4557E-02	-2,3916E-05	-3,9279E+05	1,1354E+05	-1,3251E+05

Note: *A, B, C & D* are applicable to the C_p^o/R form provided in Smith, Van Ness and Abbott (2001). The values in red, blue and black were obtained from the *Joback-Reid method*, *SVA* and *NIST*, respectively.

Joback-Reid Group Contribution Method

The Joback-Reid method is outlined in this section. Table C-3 provides the main estimation equations, which are to be used in conjunction with the data in Table C-4 and C-5 to estimate the relevant pure properties.

Table C-3. Relevant Pure Property Estimation Equations (Joback and Reid, 1987)

Property (Ideal Gas)	***Symbol***	***Unit***	***Equation***
Heat Capacity	C_p^o	$\frac{J}{mol.K}$	$C_p^o = \sum(na) - 37.93 + [\sum(nb) + 0.210]T + [\sum(nc) - 0.000391]T^2 + [\sum(nd) + 2.06 \times 10^{-7}]T^3$
Enthalpy of Formation at 298K	ΔH_{fO}^o	$\frac{kJ}{mol}$	$\Delta H_{fO}^o = 68.29 + \sum(nx)$
Gibbs Energy of Formation at 298K	ΔG_{fO}^o	$\frac{kJ}{mol}$	$\Delta G_{fO}^o = 53.88 + \sum(ny)$

Note: *a, b, c, d, x* & *y* refers to the Group contribution parameters indicated in Table C-4, and *n* refers to the respective numbers of specific groups indicated in Table C-5. The summations must be carried out for all groups present in a particular chemical species.

Table C-4. Group Contributions Parameters (Joback and Reid, 1987)

Group	**a**	**b**	**c**	**d**	**x**	**y**
-CH₃	1.95E+1	-8.08E-3	1.53E-4	-9.67E-8	-76.45	-43.96
-CH₂-	-9.09E-1	9.50E-2	-5.44E-5	1.19E-8	-20.64	8.42
=CH₂	2.36E+1	-3.81E-2	1.72E-4	-1.03E-7	-9.63	3.77
=CH-	-8.00	1.05E-1	-9.63E-5	3.56E-8	37.97	48.53

Table C-5. Number of Joback-Reid Groups in Relevant Species

Species	***Number of Groups (n)***			
	$-CH_3$	**$-CH_2-$**	**$=CH_2$**	**$=CH-$**
1-Ethene	-	-	2	-
1-Propene	1	-	1	1
1-Butene	1	1	1	1
2-Butene	2	-	-	2
1-Pentene	1	2	1	1
2-Pentene	2	1	-	2
1-Hexene	1	3	1	1
2-Hexene	2	2	-	2
3-Hexene	2	2	-	2
2-Heptene	2	3	-	2
3-Heptene	2	3	-	2
3-Octene	2	4	-	2
4-Octene	2	4	-	2
4-Nonene	2	5	-	2
5-Decene	2	6	-	2

Appendix C4: Typical Metathesis Mechanism Application & Observations

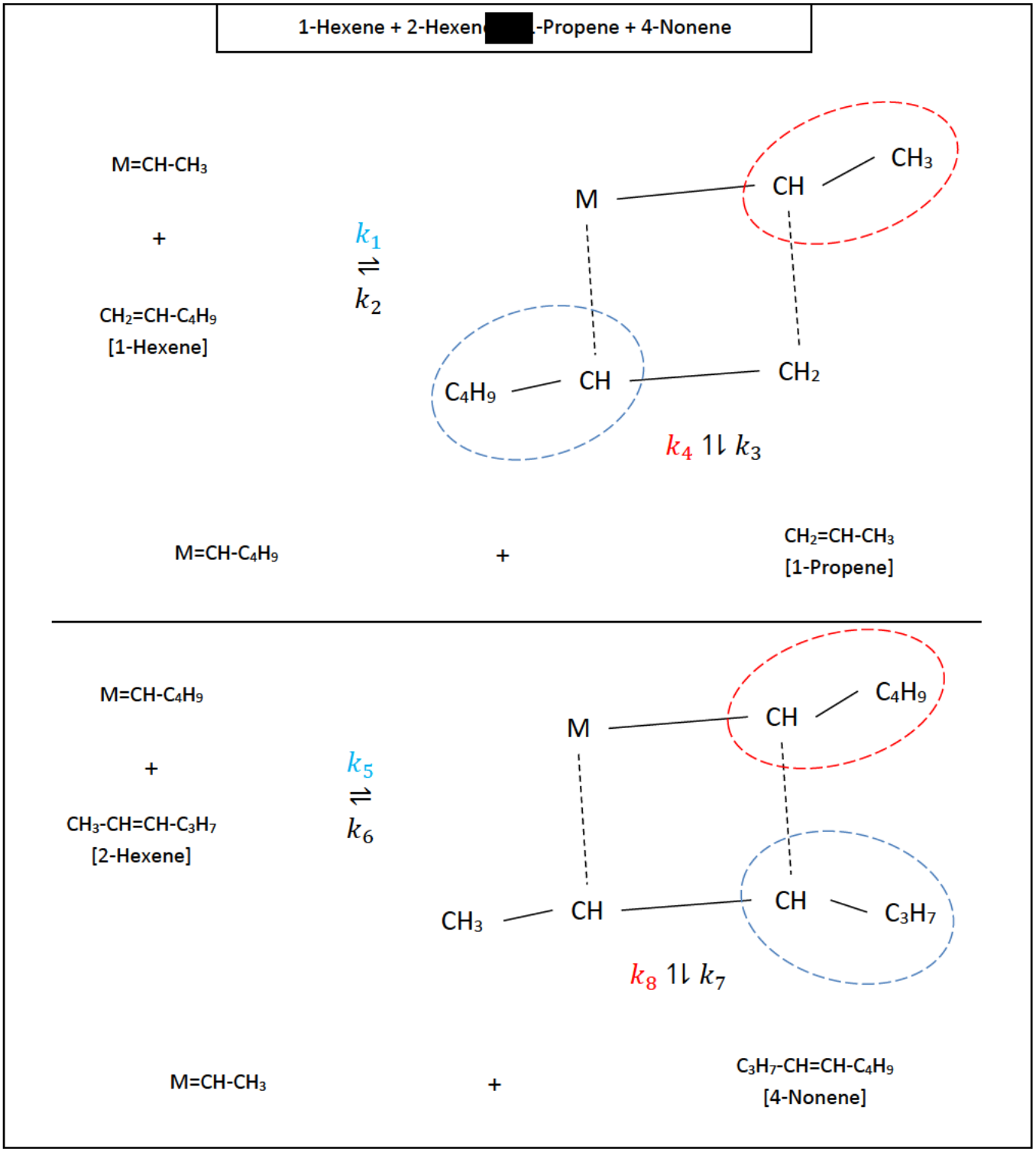


Figure C-1. Metathesis Pathway 1 for Reaction 3 of Table C-1

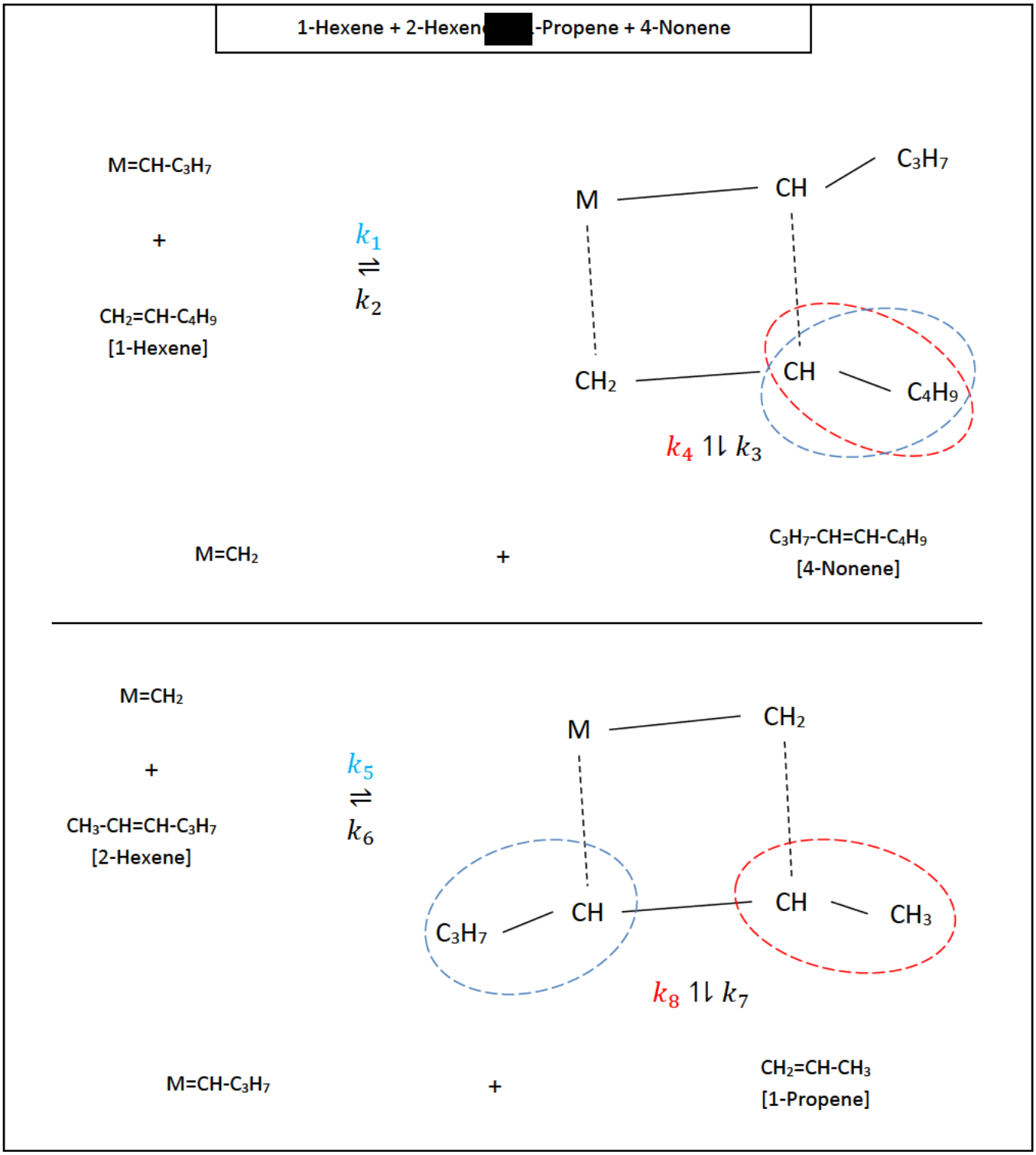


Figure C-2. Metathesis Pathway 2 for Reaction 3 of Table C-1

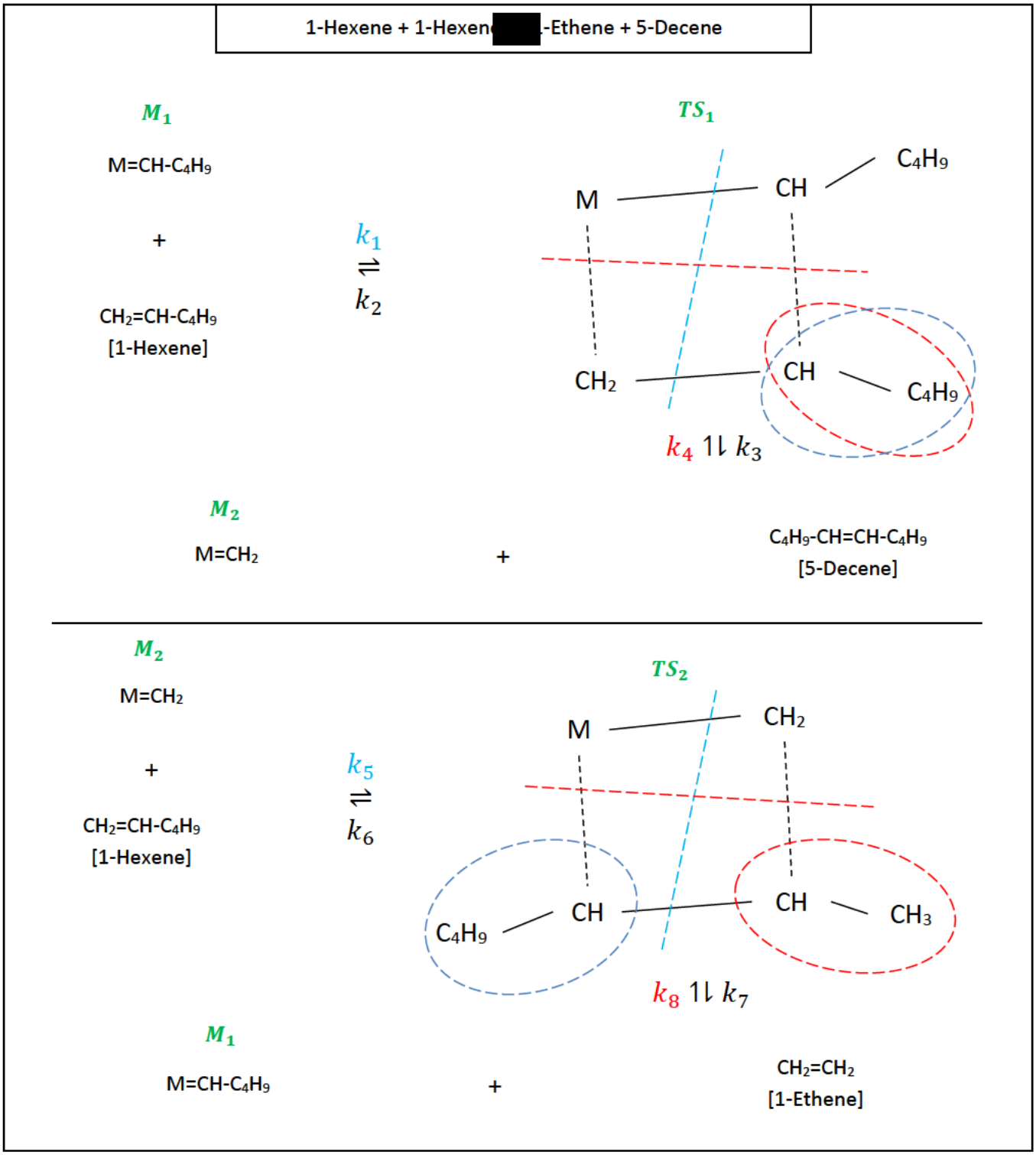


Figure C-3. Dominant Metathesis Pathway for Reaction 1 of Table C-1

Appendix C5: Quasi-Steady State Approximation Rate Expression Development

C5-1. Generalized Overall Metathesis Reaction Rate Expression

$$R_1 + R_2 \underset{k_r}{\overset{k_f}{\rightleftharpoons}} P_1 + P_2$$

Overall Reaction Rate and Equilibrium Expressions

Assuming that an elementary second order power-law relation exists,

$$r_{R_1} = k_f C_{R_1} C_{R_2} - k_r C_{P_1} C_{P_2}$$

$$r_{R_2} = k_f C_{R_1} C_{R_2} - k_r C_{P_1} C_{P_2}$$

$$r_{P_1} = k_r C_{P_1} C_{P_2} - k_f C_{R_1} C_{R_2}$$

$$r_{P_2} = k_r C_{P_1} C_{P_2} - k_f C_{R_1} C_{R_2}$$

$$C_j = \frac{F_j[mol/s]}{q_o[m^3/s]} \text{ and } K_e = \frac{k_f}{k_r}$$

Notation

R_i – *ith* Reactant
P_i – *ith* Product
r_i – Net reaction rate [mol/kg/s] of *ith* component
k_f – Forward reaction rate constant [m^6/kg/mol/s]
k_r – Reverse reaction rate constant [m^6/kg/mol /s]
F_j – *jth* component molar flowrate [mol/s]
C_j – *jth* Component concentration [mol/m^3]
K_e – Reaction equilibrium constant [-]
q_o – Bulk gas volumetric flowrate [m^3/s]

C5-2. Generalized Elementary Metathesis Reaction Steps Rate Expression

$$M_1 + R_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} TS_1$$

$$TS_1 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} M_2 + P_1$$

$$M_2 + R_2 \underset{k_6}{\overset{k_5}{\rightleftharpoons}} TS_2$$

$$TS_2 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} M_1 + P_2$$

Where, $K_A = \frac{k_1}{k_2}$ $K_B = \frac{k_3}{k_4}$ $K_C = \frac{k_5}{k_6}$ $K_D = \frac{k_7}{k_8}$

Notation

M_i – *ith* Alkylidene
TS_i – *ith* Transition state
R_i – *ith* Reactant
P_i – *ith* Product
r_i – Net reaction rate [mol/kg/s] of *ith* component
k_i – *ith* Reaction rate constant [m^6/kg/mol /s]
F_j – *jth* component molar flowrate [mol/s]
C_j – *jth* Component concentration [mol/m^3]
K_e – Reaction equilibrium constant [-]
q_o – Bulk gas volumetric flowrate [m^3/s]
$K_{A/B/C/D}$ – Elementary step equilibrium constants [-]

Quasi-Steady State Approximation

$$r_{TS_1} \approx 0 \;\rightarrow\; C_{TS_1}[-k_2 - k_3] = -k_1 C_{R_1} C_{M_1} - k_4 C_{P_1} C_{M_2}$$

$$\therefore\; C_{TS_1} = \left(k_1 C_{R_1} C_{M_1} + k_4 C_{P_1} C_{M_2}\right) \Big/ (k_2 + k_3)$$

$$r_{TS_2} \approx 0 \;\rightarrow\; C_{TS_2}[-k_6 - k_7] = -k_5 C_{R_2} C_{M_2} - k_8 C_{P_2} C_{M_1}$$

$$\therefore\; C_{TS_2} = \left(k_5 C_{R_2} C_{M_2} + k_8 C_{P_2} C_{M_1}\right) \Big/ (k_6 + k_7)$$

Quasi-Steady State Rate and Postulated Equilibrium Expression

$$r_{R_1} = k_2\left[\frac{\left(k_1 C_{R_1} C_{M_1}\right) + \left(k_4 C_{P_1} C_{M_2}\right)}{(k_2 + k_3)}\right] - \left[k_1 C_{R_1} C_{M_1}\right]$$

$$r_{R_2} = k_6\left[\frac{\left(k_5 C_{R_2} C_{M_2}\right) + \left(k_8 C_{P_2} C_{M_1}\right)}{(k_6 + k_7)}\right] - \left[k_5 C_{R_2} C_{M_2}\right]$$

$$r_{P_1} = k_3\left[\frac{\left(k_1 C_{R_1} C_{M_1}\right) + \left(k_4 C_{P_1} C_{M_2}\right)}{(k_2 + k_3)}\right] - \left[k_4 C_{P_1} C_{M_2}\right]$$

$$r_{P_2} = k_7\left[\frac{\left(k_5 C_{R_2} C_{M_2}\right) + \left(k_8 C_{P_2} C_{M_1}\right)}{(k_6 + k_7)}\right] - \left[k_8 C_{P_2} C_{M_1}\right]$$

$$C_j = \frac{F_j[mol/s]}{q_o[m^3/s]} \text{ and } K_e = \frac{forward}{reverse}$$

Lumping k_i and C_{M_i} can be performed as follows,

$$k_{1,M_1} = k_1 C_{M_1}\ ,\ k_{5,M_2} = k_5 C_{M_2}\ ,\ k_{4,M_2} = k_4 C_{M_2}\ \ \&\ k_{8,M_1} = k_8 C_{M_1}$$

This results in the following modified rate expressions,

$$r_{R_1} = k_2\left[\frac{\left(k_{1,M_1} C_{R_1}\right) + \left(k_{4,M_2} C_{P_1}\right)}{(k_2 + k_3)}\right] - \left[k_{1,M_1} C_{R_1}\right]$$

$$r_{R_2} = k_6\left[\frac{\left(k_{5,M_2} C_{R_2}\right) + \left(k_{8,M_1} C_{P_2}\right)}{(k_6 + k_7)}\right] - \left[k_{5,M_2} C_{R_2}\right]$$

$$r_{P_1} = k_3\left[\frac{\left(k_{1,M_1} C_{R_1}\right) + \left(k_{4,M_2} C_{P_1}\right)}{(k_2 + k_3)}\right] - \left[k_{4,M_2} C_{P_1}\right]$$

$$r_{P_2} = k_7\left[\frac{\left(k_{5,M_2} C_{R_2}\right) + \left(k_{8,M_1} C_{P_2}\right)}{(k_6 + k_7)}\right] - \left[k_{8,M_1} C_{P_2}\right]$$

C5-3. Modified Overall Metathesis Reaction Rate Expression

$$R_1 + R_2 \underset{k_r}{\overset{k_f}{\rightleftharpoons}} P_1 + P_2$$

With,

$$k_f = \left(k_{1,M_1} + k_{5,M_2}\right) \text{ and } k_r = \left(k_{4,M_2} + k_{8,M_1}\right)$$

Where, k_{1,M_1}, k_{5,M_2}, k_{4,M_2} and k_{8,M_1} are the *hypothesis 3* rate constants.

Modified Overall Reaction Rate and Equilibrium Expressions

Assuming that an elementary second order power-law relation exists,

$$r_{R_1} = k_f C_{R_1} C_{R_2} - k_r C_{P_1} C_{P_2}$$

$$r_{R_2} = k_f C_{R_1} C_{R_2} - k_r C_{P_1} C_{P_2}$$

$$r_{P_1} = k_r C_{P_1} C_{P_2} - k_f C_{R_1} C_{R_2}$$

$$r_{P_2} = k_r C_{P_1} C_{P_2} - k_f C_{R_1} C_{R_2}$$

$$C_j = \frac{F_j[mol/s]}{q_o[m^3/s]} \text{ and } K_e = \frac{k_f}{k_r}$$

Notation

R_i – *ith* Reactant
P_i – *ith* Product
r_i – Net reaction rate [mol/kg/s] of *ith* component
k_f – Forward reaction rate constant [m^6/kg/mol/s]
k_r – Reverse reaction rate constant [m^6/kg/mol /s]
F_j – *jth* component molar flowrate [mol/s]
C_j – *jth* Component concentration [mol/m^3]
K_e – Reaction equilibrium constant [-]
q_o – Bulk gas volumetric flowrate [m^3/s]

Table C-6. Summary of Rate constant definitions for the Generalized QSSA form

	Reaction Components				*Rate Constant Definitions*				*Alkylidenes*	
	R_1	R_2	P_1	P_2	k_1	k_4	k_5	k_8	M_1	M_2
1	1-hexene	1-hexene	5-decene	1-ethene	ch-c4h9	ch-c4h9	ch-c4h9	ch2	ch-c4h9	ch2
2	1-hexene	2-hexene	2-heptene	1-pentene	ch-c4h9	ch-c4h9	ch-c3h7	ch-c3h7	ch-ch3	ch2
3	1-hexene	2-hexene	4-nonene	1-propene	ch-c4h9	ch-c4h9	ch-c3h7	ch-ch3	ch-c3h7	ch2
4	1-hexene	3-hexene	3-octene	1-butene	ch-c4h9	ch-c4h9	ch-c2h5	ch-c2h5	ch-c2h5	ch2
5	1-hexene	1-pentene	4-nonene	1-ethene	ch-c4h9	ch-c4h9	ch-c3h7	ch2	ch-c3h7	ch2
6	1-hexene	2-heptene	5-decene	1-propene	ch-c4h9	ch-c4h9	ch-c4h9	ch-ch3	ch-c4h9	ch2
7	1-hexene	1-propene	2-heptene	1-ethene	ch-c4h9	ch-c4h9	ch-ch3	ch2	ch-ch3	ch2
8	1-hexene	4-nonene	5-decene	1-pentene	ch-c4h9	ch-c4h9	ch-c4h9	ch-c3h7	ch-c4h9	ch2
9	1-hexene	1-butene	3-octene	1-ethene	ch-c4h9	ch-c4h9	ch-c2h5	ch2	ch-c2h5	ch2
10	1-hexene	3-octene	5-decene	1-butene	ch-c4h9	ch-c4h9	ch-c4h9	ch-c2h5	ch-c4h9	ch2
11	2-hexene	3-octene	3-heptene	2-heptene	ch-c3h7	ch-c3h7	ch-c4h9	ch-c4h9	ch-c2h5	ch-ch3
12	2-hexene	3-octene	4-nonene	2-pentene	ch-c3h7	ch-c4h9	ch-c4h9	ch-c2h5	ch-c4h9	ch-ch3
13	2-hexene	2-hexene	4-octene	2-butene	ch-c3h7	ch-c3h7	ch-c3h7	ch-ch3	ch-c3h7	ch-ch3
14	2-hexene	3-hexene	3-heptene	2-pentene	ch-c3h7	ch-c3h7	ch-c2h5	ch-c2h5	ch-c2h5	ch-ch3
15	2-hexene	1-ethene	1-pentene	1-propene	ch-c3h7	ch-c3h7	ch2	ch-ch3	ch2	ch-ch3
16	2-hexene	5-decene	4-nonene	2-heptene	ch-c3h7	ch-c4h9	ch-c4h9	ch-c4h9	ch-c4h9	ch-ch3
17	2-hexene	1-pentene	1-propene	4-octene	ch-c3h7	ch-ch3	ch-c3h7	ch-c3h7	ch2	ch-c3h7
18	2-hexene	2-heptene	2-butene	4-nonene	ch-c3h7	ch-ch3	ch-c4h9	ch-c4h9	ch-ch3	ch-c3h7
19	2-hexene	1-propene	1-pentene	2-butene	ch-c3h7	ch-c3h7	ch-ch3	ch-ch3	ch2	ch-ch3
20	2-hexene	4-nonene	4-octene	2-heptene	ch-c3h7	ch-c3h7	ch-c4h9	ch-c4h9	ch-c3h7	ch-ch3
21	2-hexene	1-butene	1-pentene	2-pentene	ch-c3h7	ch-c3h7	ch-c2h5	ch-c2h5	ch2	ch-ch3
22	2-hexene	1-butene	1-propene	3-heptene	ch-c3h7	ch-ch3	ch-c2h5	ch-c3h7	ch2	ch-c3h7
23	2-hexene	3-heptene	4-octene	2-pentene	ch-c3h7	ch-c3h7	ch-c3h7	ch-c2h5	ch-c3h7	ch-ch3
24	2-hexene	2-pentene	3-heptene	2-butene	ch-c3h7	ch-c3h7	ch-c2h5	ch-ch3	ch-c2h5	ch-ch3
25	3-hexene	4-octene	3-heptene	3-heptene	ch-c2h5	ch-c3h7	ch-c3h7	ch-c3h7	ch-c3h7	ch-c2h5
26	3-hexene	1-ethene	1-butene	1-butene	ch-c2h5	ch-c2h5	ch2	ch-c2h5	ch2	ch-c2h5
27	3-hexene	5-decene	3-octene	3-octene	ch-c2h5	ch-c4h9	ch-c4h9	ch-c4h9	ch-c4h9	ch-c2h5
28	3-hexene	1-pentene	1-butene	3-heptene	ch-c2h5	ch-c2h5	ch-c3h7	ch-c3h7	ch2	ch-c2h5
29	3-hexene	2-heptene	2-pentene	3-octene	ch-c2h5	ch-c2h5	ch-c4h9	ch-c4h9	ch-ch3	ch-c2h5
30	3-hexene	1-propene	1-butene	2-pentene	ch-c2h5	ch-c2h5	ch-ch3	ch-c2h5	ch2	ch-c2h5
31	3-hexene	4-nonene	3-heptene	3-octene	ch-c2h5	ch-c3h7	ch-c4h9	ch-c4h9	ch-c3h7	ch-c2h5
32	3-hexene	2-butene	2-pentene	2-pentene	ch-c2h5	ch-c2h5	ch-ch3	ch-c2h5	ch-ch3	ch-c2h5
33	1-ethene	4-octene	1-pentene	1-pentene	ch2	ch-c3h7	ch-c3h7	ch-c3h7	ch-c3h7	ch2
34	1-ethene	3-heptene	1-butene	1-pentene	ch2	ch-c2h5	ch-c3h7	ch-c3h7	ch-c2h5	ch2
35	1-ethene	2-pentene	1-propene	1-butene	ch2	ch-ch3	ch-c2h5	ch-c2h5	ch-ch3	ch2
36	1-ethene	2-butene	1-propene	1-propene	ch2	ch-ch3	ch-ch3	ch-ch3	ch-ch3	ch2
37	5-decene	4-octene	4-nonene	4-nonene	ch-c4h9	ch-c4h9	ch-c3h7	ch-c4h9	ch-c3h7	ch-c4h9
38	5-decene	3-heptene	3-octene	4-nonene	ch-c4h9	ch-c4h9	ch-c3h7	ch-c4h9	ch-c2h5	ch-c4h9
39	5-decene	2-pentene	2-heptene	3-octene	ch-c4h9	ch-c4h9	ch-c2h5	ch-c4h9	ch-ch3	ch-c4h9
40	5-decene	2-butene	2-heptene	2-heptene	ch-c4h9	ch-c4h9	ch-ch3	ch-c4h9	ch-ch3	ch-c4h9
41	1-pentene	2-heptene	4-nonene	1-propene	ch-c3h7	ch-c4h9	ch-c4h9	ch-ch3	ch-c4h9	ch2
42	1-pentene	4-nonene	4-octene	1-hexene	ch-c3h7	ch-c3h7	ch-c4h9	ch-c4h9	ch-c3h7	ch2
43	1-pentene	3-octene	3-heptene	1-hexene	ch-c3h7	ch-c3h7	ch-c4h9	ch-c4h9	ch-c2h5	ch2
44	1-pentene	3-octene	4-nonene	1-butene	ch-c3h7	ch-c4h9	ch-c4h9	ch-c2h5	ch-c4h9	ch2
45	1-pentene	3-heptene	4-octene	1-butene	ch-c3h7	ch-c3h7	ch-c3h7	ch-c2h5	ch-c3h7	ch2
46	2-heptene	1-propene	1-hexene	2-butene	ch-c4h9	ch-c4h9	ch-ch3	ch-ch3	ch2	ch-ch3
47	2-heptene	1-butene	1-hexene	2-pentene	ch-c4h9	ch-c4h9	ch-c2h5	ch-c2h5	ch2	ch-ch3
48	2-heptene	1-butene	3-octene	1-propene	ch-c4h9	ch-c4h9	ch-c2h5	ch-ch3	ch-c2h5	ch-ch3
49	2-heptene	3-heptene	4-nonene	2-pentene	ch-c4h9	ch-c4h9	ch-c3h7	ch-c2h5	ch-c3h7	ch-ch3
50	2-heptene	2-pentene	3-octene	2-butene	ch-c4h9	ch-c4h9	ch-c2h5	ch-ch3	ch-c2h5	ch-ch3
51	1-propene	3-octene	2-pentene	1-hexene	ch-ch3	ch-c2h5	ch-c4h9	ch-c4h9	ch-c2h5	ch2
52	1-propene	3-heptene	2-pentene	1-pentene	ch-ch3	ch-c2h5	ch-c3h7	ch-c3h7	ch-c2h5	ch2
53	1-propene	2-pentene	2-butene	1-butene	ch-ch3	ch-ch3	ch-c2h5	ch-c2h5	ch-ch3	ch2
54	4-nonene	1-butene	1-hexene	3-heptene	ch-c4h9	ch-c4h9	ch-c2h5	ch-c3h7	ch2	ch-c3h7
55	4-nonene	3-heptene	3-octene	4-octene	ch-c4h9	ch-c4h9	ch-c3h7	ch-c3h7	ch-c2h5	ch-c3h7

Appendix D1: Algorithms

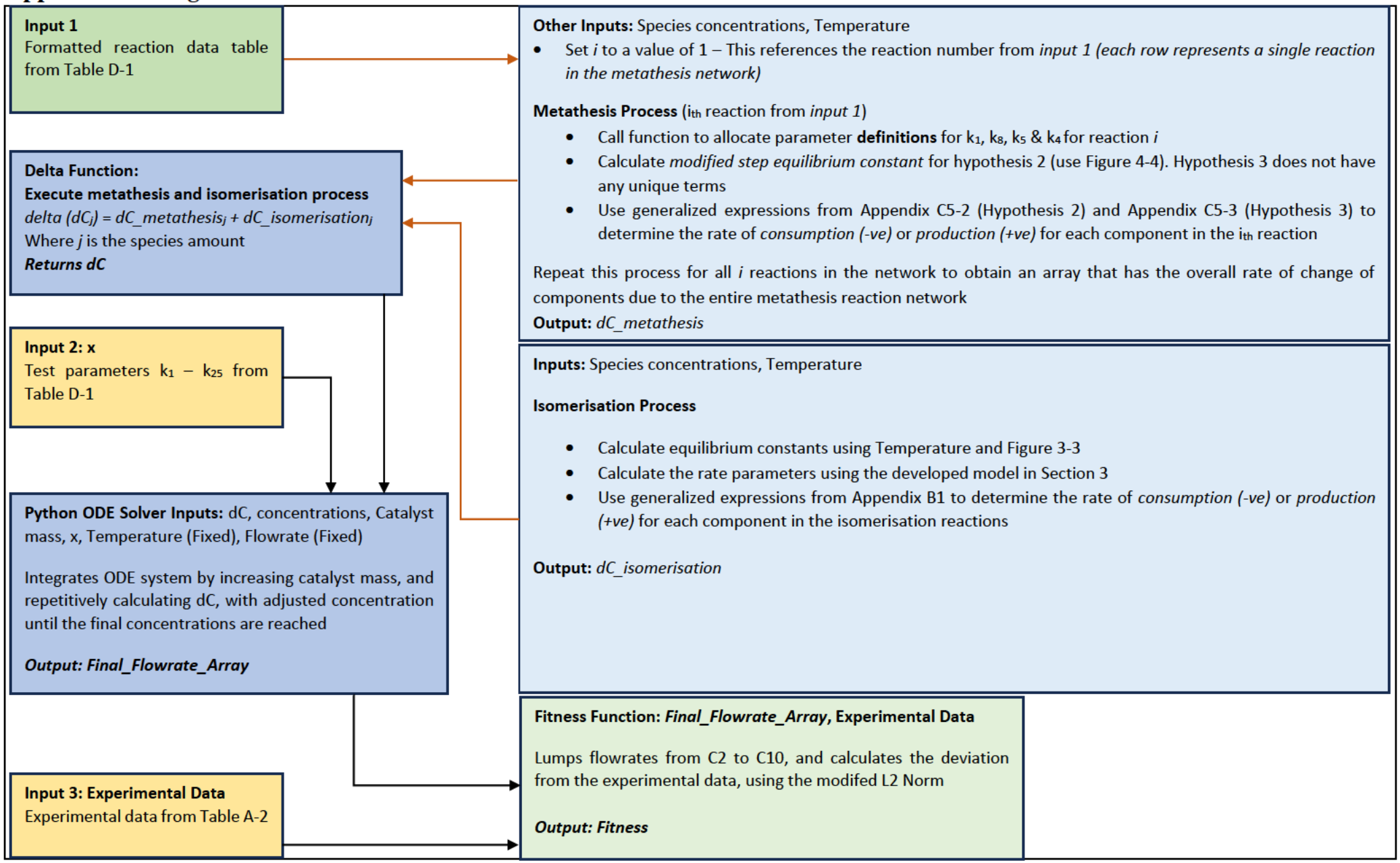


Figure D-1. Reacting System Model – Hypothesis Validation

Inputs: Np, D, ub, lb, num_iter, w, c1, c2

PSO Process

- Generate a (seeded) random input population (P), with D rows and Np columns
- Set i to zero, and initialize blank fitness vector (f) of size Np
- Repeat until i = Np:
  - Execute F1 with P at column i as input
  - Store outputted fitness at position i of vector f
  - i = i +1
- Set Pbest to P, Pbest_fitness to f, gbest_fitness to lowest value in Pbest_fitness, and gbest to parameters that yield lowest Pbest_fitness
- Set r1 and r2 to random value arrays, with D rows and Np columns
- Set t and i to zero
- Repeat until t = num_iter:
  - Repeat until i = Np:
    - Calculate velocity at future state (2) for parameter set i (X_i):
      - $v_i|_2 = w.v_i|_1 + c_1.r_1.\left(P_{best,i} - X_i|_1\right) + c_2.r_2.(g_{best} - X_i|_1)$
    - Calculate parameter at future state (2) for parameter set i (X_i):
      - $X_i|_2 = X_i|_1 + v_i|_2$
    - Bound $X_i|_2$ using ub and lb
    - Execute F1 with $X_i|_2$ as input
    - Update fitness at position i for vector f
    - Update $X_i|_1$ with $X_i|_2$
    - IF: f at position i < $P_{best\_fitness,i}$:
      - Set $P_{best,i}$ to $X_i|_2$, and $P_{best\_fitness,i}$ to f at position i
    - IF: $P_{best\_fitness,i}$ < $g_{best\_fitness}$:
      - Set $g_{best,}$ to $P_{best,i}$, and $g_{best\_fitness}$ to $P_{best\_fitness,i}$
    - i = i +1
  - t = t + 1
- Output g_{best} and $g_{best\_fitness}$ as best parameter set and fitness

Function: F1

Variable Input:
Parameters $k_1 - k_{25}$

Reacting System Model:

The entire reacting system model described in Figure D-1, excluding the parameter input

Output: Fitness

Notation:

- **Np - Number of parameter sets**
- **D - Number of parameters (Fixed at 25)**
- **ub - Upper bound**
- **lb - Lower bound**
- **num_iter – Number of iterations**
- **w - Inertia**
- **c1 - Constant**
- **c2 - Constant**

Figure D-2. PSO – Reacting System Parameter Search

Inputs: Np, dvs, iter_count, pc, nc, lb, ub, pm, nm

GA Process

- Generate a (seeded) random input population (P), with Np rows and dvs columns
- Initialize population fitness (fp) and offspring fitness (fo) to zero vector with size Np rows and 1 column
- Set i to zero , j to zero
- Repeat until i = Np:
  - Execute F1 with input parameter set i from array P
  - Store fitness at position i in fp
  - i = i + 1
- Repeat until j = iter_count:
  - Update P, fp using ***tournament selection***, with P and fp as inputs
  - Generate Offspring (O) using ***crossover***, with pc and nc as inputs
  - Update Offspring (O) using ***mutation***, with pm and nm as inputs
  - Calculate fitness fo, for offspring (O), as done for fp
  - Determine updated P, fp using ***survival*** to select the top Np parameters from all of O and P, when ordered by fitness
  - j = j + 1
- Output P and fp at position zero, as best parameter set and fitness

Mutation (Polynomial): O, Np, dvs, pm, nm, ub, lb

- Set i to zero
- Repeat until i = Np:
  - Set u to a random number (0 to 1)
  - IF: u<pm:
    - Set r to a random array (0 to 1) and delta to a zero array with 1 row and dvs columns, Set j to zero
    - Repeat until j = dvs:
      - IF: $r|_j \geq 0.5$
        - $delta|_j = 1 - \left(2.\left(1 - r|_j\right)\right)^{\left(\frac{1}{n_m+1}\right)}$
      - ELSE:
        - $delta|_j = \left(2.r|_j\right)^{\left(\frac{1}{n_m+1}\right)} - 1$
      - j = j + 1
    - $O1 = O|_i + (ub - lb).delta$
    - Bound and update $O|_i$ with O1
  - i = i +1
- Return updated O

Function: F1

Variable Input:
Parameters $k_1 - k_{25}$

Reacting System Model:

The entire reacting system model described in Figure D-1, excluding the parameter input

Output: Fitness

Notation:

- **Np - Number of parameter sets**
- **dvs - Number of parameters (Fixed at 25)**
- **ub - Upper bound, lb - Lower bound**
- **iter_count – Number of iterations**
- **pc – crossover probability, nc - damper**
- **pm – mutation probability, nm - damper**

Tournament Selection: P, fp

- Initialise pool and pool_fitness to zero arrays of size P and fp, respectively
- Set count, i, j to zero
- Repeat until j = 2:
  - Set T to P
  - Repeat until i = (Np/2):
    - Set T1 and f1 to random parameter set, and fitness, A, from T, fp
    - Remove set A data from T, fp
    - Set T2 and f2 to random parameter set, and fitness, B, from T, fp
    - IF: f1<f2
      - Store T1, f1 in pool, pool_fitness at position count
      - count = count + 1
    - ELSE:
      - Store T2, f2 in pool, pool_fitness at position count
      - count = count + 1
    - i = i + 1
  - j = j + 1
- Update P, fp with pool, pool_fitness
- Return P, fp

Crossover (SBX): P, fp, Np, dvs, pc, nc, ub, lb

- Initialise zero array Offspring (O) with size P, Set T to P, count to zero, i to zero
- Repeat until i = (Np/2):
  - Set T1 and f1 to random parent set, and fitness, A, from T, fp
  - Remove set A data from T, fp
  - Set T2 and f2 to random parent set, and fitness, B, from T, fp
  - Set r to a random number (0 to 1)
  - IF: r<pc
    - Set u to a random array (0 to 1) and beta to a zero array with 1 row and dvs columns, Set j to zero
    - Repeat until j = dvs:
      - IF: $u|_j > 0.5$
        - $beta|_j = \frac{1}{2.(1-r|_j)}^{\frac{1}{n_c+1}}$
      - ELSE:
        - $beta|_j = \left(2.u|_j\right)^{\left(\frac{1}{n_c+1}\right)}$
      - j = j + 1
    - $O1 = 0{,}5.\left((1 + beta).T1 + (1 - beta).T2\right)$ & $O2 = 0{,}5.\left((1 - beta).T1 + (1 + beta).T2\right)$
    - Store O1 in $O|_{count}$, Increment count, Store O2 in $O|_{count}$, Increment count
  - ELSE:
    - Store T1 in $O|_{count}$, Increment count, Store T2 in $O|_{count}$, Increment count
  - i = i + 1
- Return generated O

Figure D-3. GA - Reacting System Parameter Search

Appendix D2: Randomized Hypothesis Testing

Table D-1. Reaction Model Simulation Metathesis Data

Rxn	r_1	r_2	p_1	p_2	k_1	k_4	k_5	k_8	m_1	m_2	Keq (693,15K)	Keq (713,15K)	Keq (733,15K)
1	1	1	5	4	A	A	A	E	A	E	0,09037664	0,09285320	0,09529384
2	1	2	7	6	A	A	B	B	D	E	0,70919111	0,70455210	0,70022443
3	1	2	9	8	A	A	B	D	B	E	0,93854309	0,94382121	0,94891557
4	1	3	11	10	A	A	C	C	C	E	0,76393718	0,76242525	0,76104064
5	1	6	9	4	A	A	B	E	B	E	0,08903091	0,09137654	0,09368312
6	1	7	5	8	A	A	A	D	A	E	1,34340296	1,36125293	1,37845882
7	1	8	7	4	A	A	D	E	D	E	0,06727441	0,06821157	0,06913071
8	1	9	5	6	A	A	A	B	A	E	1,01511528	1,01616027	1,01719328
9	1	10	11	4	A	A	C	E	C	E	0,10188351	0,10455178	0,10717418
10	1	11	5	10	A	A	A	C	A	E	0,88705855	0,88810739	0,88914926
11	2	11	15	7	B	B	A	A	C	D	0,35785770	0,35015728	0,34302943
12	2	11	9	14	B	A	A	C	A	D	0,65507459	0,64923484	0,64376158
13	2	2	13	12	B	B	B	D	B	D	0,99312932	0,99298456	0,99284773
14	2	3	15	14	B	B	C	C	C	D	0,28897386	0,28147891	0,27457174
15	2	4	6	8	B	B	E	D	E	D	10,54176666	10,32892314	10,12899214
16	2	5	9	7	B	A	A	A	A	D	0,69863110	0,69334742	0,68838877
17	2	6	8	13	B	D	B	B	E	B	0,92457242	0,92881622	0,93288164
18	2	7	12	9	B	D	A	A	D	B	1,42152927	1,43215269	1,44226947
19	2	8	6	12	B	B	D	D	E	D	1,07414985	1,06908615	1,06428048
20	2	9	13	7	B	B	A	A	B	D	0,69863445	0,69335104	0,68839267
21	2	10	6	14	B	B	C	C	E	D	0,74964187	0,74284558	0,73646797
22	2	10	8	15	B	D	C	B	E	B	0,54195643	0,53670606	0,53180266
23	2	15	13	14	B	B	B	C	B	D	1,27888174	1,28555848	1,29190301
24	2	14	15	12	B	B	C	D	C	D	0,77656072	0,77241492	0,76851569
25	3	13	15	15	C	B	B	B	B	C	0,22595823	0,21895457	0,21253278
26	3	4	10	10	C	C	E	C	E	C	7,49814344	7,29232236	7,10097031
27	3	5	11	11	C	A	A	A	A	C	0,86120265	0,85848318	0,85592000
28	3	6	10	15	C	C	B	B	E	C	0,38548255	0,37891981	0,37282237
29	3	7	14	11	C	C	A	A	D	C	0,80751053	0,80386423	0,80043202
30	3	8	10	14	C	C	D	C	E	C	0,53320496	0,52445636	0,51630381
31	3	9	15	11	C	B	A	A	B	C	0,44113123	0,43355484	0,42651153
32	3	12	14	14	C	C	D	C	D	C	0,37212011	0,36441412	0,35727538
33	4	13	6	6	E	B	B	B	B	E	11,40177500	11,12052411	10,85774623
34	4	15	10	6	E	C	B	B	C	E	19,45131745	19,24502803	19,04652417
35	4	14	8	10	E	D	C	C	D	E	14,06240385	13,90453596	13,75347271
36	4	12	8	8	E	D	D	D	D	E	9,81405585	9,66145068	9,51722059
37	5	13	9	9	A	A	B	A	B	A	0,99999521	0,99999478	0,99999433
38	5	15	11	9	A	A	B	A	C	A	1,95225955	1,98010286	2,00679217
39	5	14	7	11	A	A	C	A	D	A	1,06649093	1,06794549	1,06932255
40	5	12	7	7	A	A	D	A	D	A	0,49146445	0,48412954	0,47729553
41	6	7	9	8	B	A	A	D	A	E	1,32339940	1,33960456	1,35515918
42	6	9	13	1	B	B	A	A	B	E	0,98511451	0,98410187	0,98310290
43	6	11	15	1	B	B	A	A	C	E	0,50459980	0,49699273	0,48988497
44	6	11	9	10	B	A	A	C	A	E	0,87385006	0,87398358	0,87412026
45	6	15	13	10	B	B	B	C	B	E	1,70599029	1,73058642	1,75418763
46	7	8	1	12	A	A	D	D	E	D	1,51461268	1,51739829	1,51991338
47	7	10	1	14	A	A	C	C	E	D	1,05703787	1,05435153	1,05175989
48	7	10	11	8	A	A	C	D	C	D	1,51444678	1,53275713	1,55031206
49	7	15	9	14	A	A	B	C	B	D	1,83054492	1,85412353	1,87669490
50	7	14	11	12	A	A	C	D	C	D	2,17002659	2,20590854	2,24037831
51	8	11	14	1	D	C	A	A	C	E	0,69796964	0,68787906	0,67841818
52	8	15	14	6	D	C	B	B	C	E	1,38321425	1,38408273	1,38485200
53	8	14	12	10	D	D	C	C	D	E	1,43288402	1,43917683	1,44511442
54	9	10	1	15	A	A	C	B	E	B	0,57744438	0,56865225	0,56043201
55	9	15	11	13	A	A	B	B	C	B	1,95226890	1,98011320	2,00680354

Table D-2. Randomly Selected Parameter Sets

Defined Parameter $\left[\frac{m^3}{sec.kg}\right]$	In Simulation	Set 1	Set 2	Set 3	Set 4
kaa	k1	1,00E-05	1,00E-01	1,00E-05	1,00E-09
kab	k2	1,00E-06	2,00E-01	2,00E-05	2,00E-09
kac	k3	1,00E-07	3,00E-01	3,00E-05	3,00E-09
kad	k4	1,00E-08	4,00E-01	4,00E-05	4,00E-09
kae	k5	1,00E-09	5,00E-01	5,00E-05	5,00E-09
kba	k6	2,00E-05	1,00E-02	1,00E-06	1,00E-10
kbb	k7	2,00E-06	2,00E-02	2,00E-06	2,00E-10
kbc	k8	2,00E-07	3,00E-02	3,00E-06	3,00E-10
kbd	k9	2,00E-08	4,00E-02	4,00E-06	4,00E-10
kbe	k10	2,00E-09	5,00E-02	5,00E-06	5,00E-10
kca	k11	3,00E-05	1,00E-03	1,00E-07	1,00E-11
kcb	k12	3,00E-06	2,00E-03	2,00E-07	2,00E-11
kcc	k13	3,00E-07	3,00E-03	3,00E-07	3,00E-11
kcd	k14	3,00E-08	4,00E-03	4,00E-07	4,00E-11
kce	k15	3,00E-09	5,00E-03	5,00E-07	5,00E-11
kda	k16	4,00E-05	1,00E-04	1,00E-08	1,00E-12
kdb	k17	4,00E-06	2,00E-04	2,00E-08	2,00E-12
kdc	k18	4,00E-07	3,00E-04	3,00E-08	3,00E-12
kdd	k19	4,00E-08	4,00E-04	4,00E-08	4,00E-12
kde	k20	4,00E-09	5,00E-04	5,00E-08	5,00E-12
kea	k21	5,00E-05	1,00E-05	1,00E-09	1,00E-13
keb	k22	5,00E-06	2,00E-05	2,00E-09	2,00E-13
kec	k23	5,00E-07	3,00E-05	3,00E-09	3,00E-13
ked	k24	5,00E-08	4,00E-05	4,00E-09	4,00E-13
kee	k25	5,00E-09	5,00E-05	5,00E-09	5,00E-13

Note: a refers to the ch-c_4h_9 group, b refers to the ch-c_3h_7 group, c refers to the ch-c_2h_5 group, d refers to the ch-ch_3 group, and e refers to the ch_2 group.

Table D-3. Hypothesis [Model 2 & 3] Results

Outputs [mol/s]	**Set 1**		**Set 2**		**Set 3**		**Set 4**		**Experimental Data**
	Model 2	**Model 3**	**Model 2**	**Model 3**	**Model 2**	**Model 3**	**Model 2**	**Model 3**	
C2	3,60819E-13	-2,3E-07	7,38E-06	0	1,8E-08	-1,4E-06	1,82E-12	-1,4E-10	2,6312E-07
C3	1,28595E-11	-1,9E-08	2,46E-05	0	3,36E-08	-5,3E-07	3,23E-12	-4,7E-11	3,1632E-07
C4	1,11961E-10	-5,5E-08	1,76E-05	0	2,41E-08	-3,1E-07	2,23E-12	-2,8E-11	1,7025E-07
C5	4,98981E-11	-1,4E-09	1,35E-06	0	5,8E-08	-9,8E-07	5,62E-12	-8,6E-11	9,8534E-07
C6	5,1443E-05	5,22E-05	4,5E-07	0	5,08E-05	5,81E-05	5,16E-05	5,16E-05	4,6818E-05
C7	2,40608E-09	2,23E-09	6,42E-08	0	2,44E-07	-9,3E-07	2,53E-11	-8,6E-11	6,8505E-07
C8	2,87905E-08	-5,3E-08	1,14E-07	0	2,41E-07	-2,9E-07	2,49E-11	-2,8E-11	3,2143E-07
C9	8,65835E-08	-1,5E-08	4,7E-09	0	1,48E-07	-4,8E-07	1,54E-11	-4,7E-11	3,6396E-07
C10	7,14857E-08	-2,3E-07	1,78E-09	0	6,99E-08	-1,5E-06	7,15E-12	-1,4E-10	3,4292E-07
C11	0	0	0	0	0	0	0	0	0
C12	0	0	0	0	0	0	0	0	0
C13	0	0	0	0	0	0	0	0	0
C14	0	0	0	0	0	0	0	0	0
C15	0	0	0	0	0	0	0	0	0
C16	0	0	0	0	0	0	0	0	0

Note: Model 2 is based on Hypothesis 2, and Model 3 is based on Hypothesis 3. Experimental data was obtained from Table A-2, Run 1 Set 1. C2 implies carbon chain of length 2, which extends for C3 onwards.

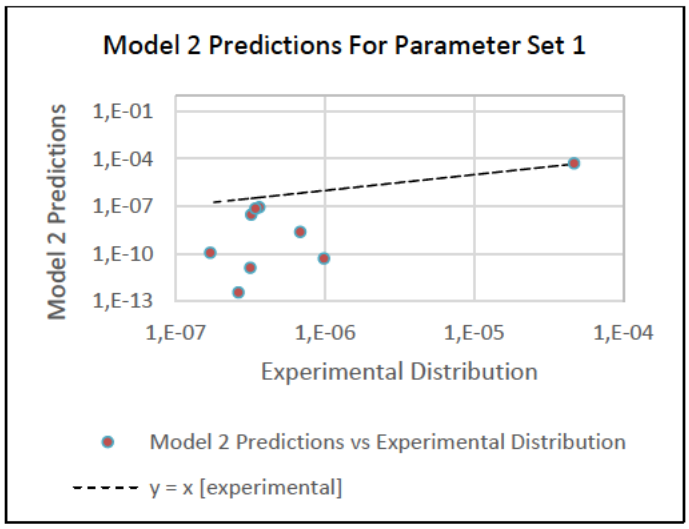


Figure D-4. Model 2 Results with Parameter Set 1

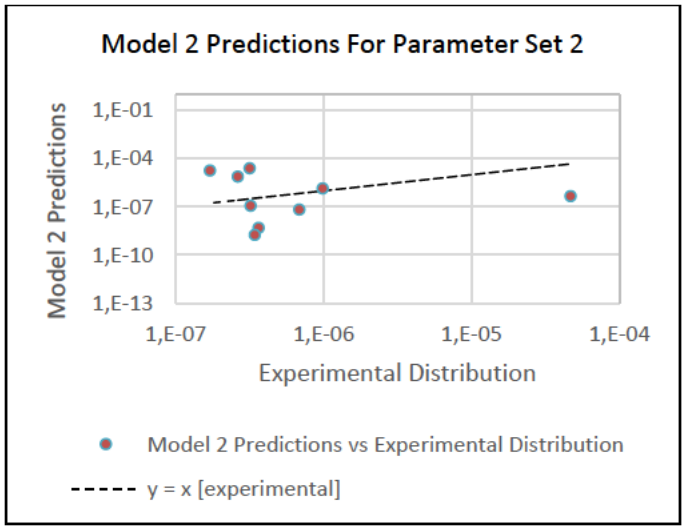


Figure D-5. Model 2 Results with Parameter Set 2

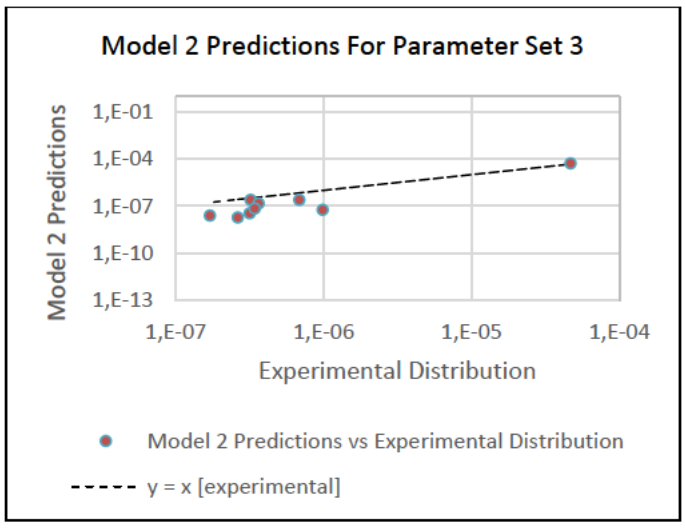


Figure D-6. Model 2 Results with Parameter Set 3

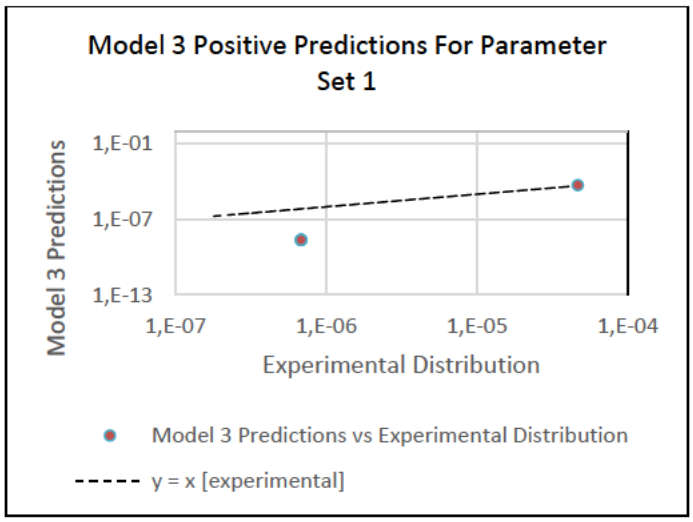


Figure D-7. Model 3 Positive Results with Parameter Set 1

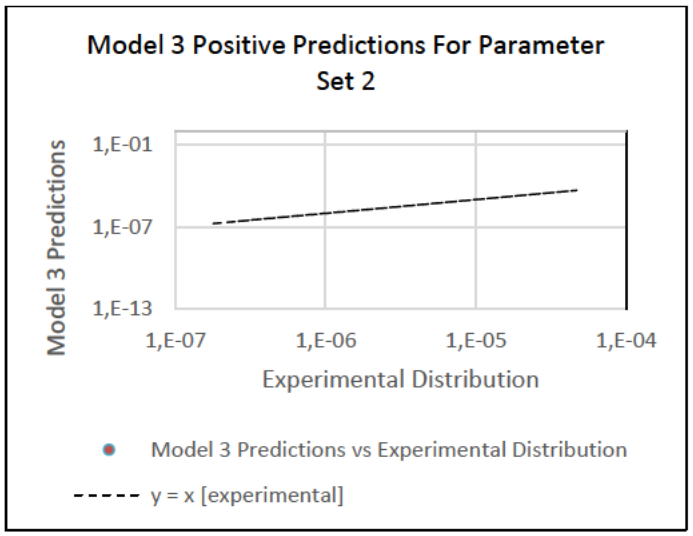


Figure D-8. Model 3 Positive Results with Parameter Set 2

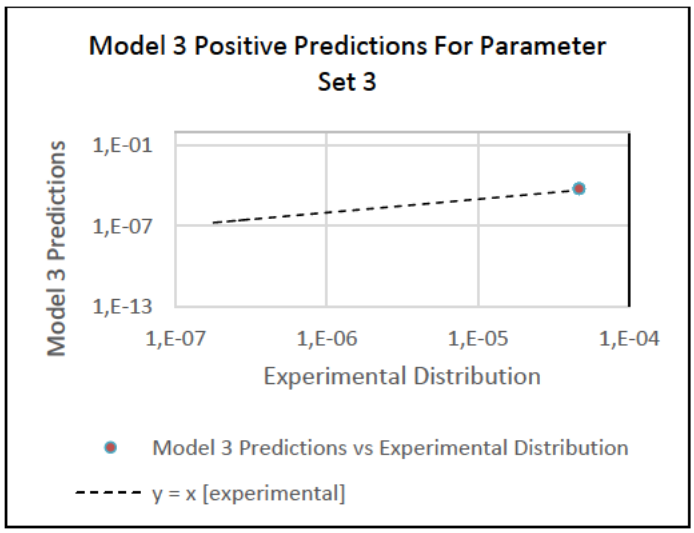


Figure D-9. Model 3 Positive Results with Parameter Set 3

Appendix D3: Optimized Arrhenius Constants

Table D-4. Optimized Parameters and Final Step optimization Parameters

Parameter $\left[\frac{m^3}{sec.kg}\right]$	**Run 1 - Set 1**	**Run 1 - Set 2**	**Run 2 - Set 1**	**Run 2 - Set 2**	**Run 3 - Set 1**	**Run 3 - Set 2**	**Run 4 - Set 1**	**Run 4 - Set 2**	**Run 5 - Set 1**	**Run 5 - Set 2**
k1	4,762E-05	1,135E-04	8,760E-04	6,813E-04	1,838E-04	5,052E-04	8,999E-04	9,328E-04	6,561E-04	3,049E-04
k2	2,367E-09	4,345E-05	1,967E-04	1,519E-04	1,500E-09	4,368E-04	4,515E-06	1,154E-04	1,259E-09	1,292E-04
k3	1,814E-09	1,099E-05	1,867E-04	1,373E-03	2,120E-04	1,179E-07	1,929E-04	3,254E-04	6,821E-06	4,978E-04
k4	4,371E-05	9,108E-10	1,285E-09	6,272E-04	1,978E-04	7,260E-04	1,888E-04	5,367E-05	5,048E-04	4,438E-04
k5	2,214E-04	2,103E-04	1,036E-03	6,669E-04	4,545E-04	9,682E-04	3,345E-04	2,363E-04	7,169E-04	9,101E-04
k6	4,382E-04	5,155E-08	7,041E-04	2,177E-03	2,386E-04	2,307E-05	1,000E-03	7,577E-04	8,006E-04	7,195E-04
k7	2,882E-04	1,729E-05	3,119E-04	6,593E-04	1,624E-04	3,689E-04	1,096E-04	5,758E-05	4,251E-04	1,000E-09
k8	2,988E-05	1,388E-04	1,149E-03	8,018E-04	1,781E-04	7,711E-04	3,337E-04	4,255E-04	1,718E-04	5,672E-04
k9	4,751E-05	8,762E-06	9,846E-04	5,061E-04	2,036E-04	9,993E-04	1,051E-04	3,527E-04	4,489E-05	3,473E-04
k10	1,202E-04	1,020E-04	9,201E-04	8,460E-04	5,084E-04	3,939E-04	4,838E-04	7,321E-04	9,779E-04	7,081E-04
k11	5,329E-08	5,320E-05	1,117E-03	1,875E-04	6,768E-04	3,430E-04	2,155E-05	9,192E-04	8,840E-07	9,554E-04
k12	3,789E-05	3,351E-05	5,524E-05	5,180E-05	1,762E-04	2,191E-04	3,949E-04	8,646E-04	7,724E-04	1,000E-03
k13	4,744E-06	9,042E-10	5,060E-04	6,110E-05	7,572E-06	1,000E-09	2,131E-04	7,206E-04	1,891E-05	4,882E-04
k14	6,289E-10	9,314E-04	3,019E-04	4,183E-09	2,490E-04	3,621E-04	3,711E-08	9,580E-10	8,666E-04	8,734E-04
k15	2,773E-05	1,302E-04	8,557E-04	1,808E-03	6,230E-04	3,267E-06	6,981E-04	9,061E-04	2,608E-09	1,000E-03
k16	3,560E-07	6,161E-05	1,936E-04	5,825E-04	1,144E-09	7,372E-04	7,395E-04	6,152E-04	4,989E-04	7,028E-04
k17	2,363E-09	7,022E-05	1,305E-03	8,710E-04	1,412E-09	5,841E-06	8,182E-04	1,359E-04	6,334E-04	1,571E-04
k18	3,956E-05	3,613E-05	4,633E-04	1,793E-04	3,093E-05	1,799E-04	5,359E-04	5,177E-04	4,905E-04	2,984E-04
k19	1,569E-09	3,109E-04	4,908E-04	5,105E-04	1,257E-04	3,687E-07	4,155E-05	9,441E-04	3,224E-04	4,789E-04
k20	2,049E-09	2,666E-04	7,968E-04	7,904E-04	1,549E-09	3,866E-04	6,668E-04	3,066E-04	5,802E-05	2,867E-04
k21	4,998E-05	4,202E-05	4,654E-04	2,610E-04	1,398E-04	4,958E-04	6,870E-04	5,859E-04	3,628E-04	9,793E-04
k22	3,525E-05	8,865E-05	8,935E-04	6,587E-04	4,455E-04	9,962E-04	5,806E-04	5,385E-04	2,427E-05	3,001E-04
k23	5,211E-06	9,570E-06	6,004E-04	1,613E-09	2,293E-05	2,576E-05	8,788E-06	5,802E-04	4,645E-05	7,126E-06
k24	1,589E-04	4,001E-04	2,740E-04	2,891E-04	3,052E-06	2,864E-04	1,503E-04	5,193E-04	4,083E-05	7,661E-04
k25	3,673E-07	2,743E-07	1,693E-04	5,994E-04	7,798E-04	7,678E-04	8,462E-08	6,643E-04	4,387E-04	4,850E-05
Final Optimisation Hyperparameters										
Method	PSO	PSO	PSO	PSO	PSO	PSO	PSO	PSO	PSO	GA
Final Fitness	1,305	1,584	1,255	2,287	0,626	0,568	1,363	0,935	1,202	1,384
Np	50	60	60	60	50	50	60	50	50	50
Iterations	20	25	20	30	25	30	25	25	35	50
c1	0,01	0,01	0,02	0,01	0,01	0,01	0,02	0,02	0,02	-
c2	0,08	0,05	0,06	0,05	0,05	0,05	0,05	0,06	0,04	-
Search Space	1	1	1	1	1	1	1	1	1	-
T [°C]	420	420	460	460	420	420	460	460	440	440

Appendix E1: 1-Hexene Isomerisation [Matlab]

```
%% Script: Regression of Isomerisation Kinetic parameters @473K
%Shodhanga Run 1 Dataset
clc;clear all;
%% User Defined Parameters:
T=473; %[K]
Po=98.8; %[kPa] - Pressure at  reactor entrance
P=98.8; %[kPa]
Faog=0.2607267*(10^-3)/(60); % [kg/s] - From data
Fao=Faog/(84.161); %[kmol/s]
Fn2=3.451*(10^-8); %[kmol/s] - Excluded Now
qo=(Fao)*8.314*T/P; %[m^3/s] - See data
wcat=5.214534e-3; %[kg] - Catalyst mass
%% Predicted Final Amounts (wt%):
    wa1=0.82;
    wb1=0.17;
    wc1=0.01;
    Fa=wa1*Faog;% [kg/s]
    Fb=wb1*Faog;% [kg/s]
    Fc=wc1*Faog;% [kg/s]
%% Processing:
%       Guess constituent formation rate constants:
    v1=linspace(6.5e-5,8e-5,50);% [1 to 2-Hex]
    v2=linspace(3.5e-6,5e-6,50);%[1 to 3-Hex]
    v3=linspace(2e-6,3e-6,50);% [2 to 3-Hex]
    K1e=7.1455;%[1 to 2-Hex] - equilibrium constant
    K2e=3.4336;%[1 to 3-Hex] - equilibrium constant
    K3e=0.2309;%[2 to 3-Hex] - equilibrium constant
%       Solving Ode System:
    k1t=0;% True k1
    k2t=0;% True k2
    k3t=0;% True k3
    wat=0;% True 1-Hex Fraction
    wbt=0;% True 2-Hex Fraction
    wct=0;% True 3-Hex Fraction
    k=[];
    w_frac=[];
    eo=[];
    s1=3;
    lowsum_pos=0;
    j=1;
for a=1:1:length(v1)
    for b=1:1:length(v2)
        for c=1:1:length(v3)
    %       Parameter assumptions & Call solver
        k1=v1(1,a);
        k2=v2(1,b);
        k3=v3(1,c);
        Ao=[Fao,zeros(1,2)]; % Initial Values for all 15 components
        options=odeset('NonNegative',[1 2 3]);% 3 Unique Species
    [w,F]=ode23s(@(w,F) solver_iso(w,F,k1,k2,k3,K1e,K2e,K3e,qo,Po,P),[0
wcat],Ao,options);
    % Mass Flows
        L=size(F);
        Lr=L(1,1);
        Fai=F(Lr,1)*84.161;% [kg/s]
        Fbi=F(Lr,2)*84.161;% [kg/s]
        Fci=F(Lr,3)*84.161;% [kg/s]
        Fi_tot=Fai+Fbi+Fci;% [kg/s]
    % Mass Fractions and Target Error:
        wa2=(Fai/Fi_tot);
        wb2=(Fbi/Fi_tot);
        wc2=(Fci/Fi_tot);
        e1=abs((wa1-wa2)/wa1);
```

```
        e2=abs((wb1-wb2)/wb1);
        e3=abs((wc1-wc2)/wc1);
            s2=e1+e2+e3;
%       Storing values:
        k(j,1)=k1;
        k(j,2)=k2;
        k(j,3)=k3;
        eo(j,1)=e1;
        eo(j,2)=e2;
        eo(j,3)=e3;
        eo(j,4)=e1+e2+e3;
        w_frac(j,1)=wa2;
        w_frac(j,2)=wb2;
        w_frac(j,3)=wc2;
            if s2<s1
                lowsum_pos=j;
                s1=s2;
            end
%       Increment counter:
        j=j+1;
        end
    end
end
beep
eo(lowsum_pos,[1,2,3,4])
    k1t=k(lowsum_pos,1)
    k2t=k(lowsum_pos,2)
    k3t=k(lowsum_pos,3)
        wat=w_frac(lowsum_pos,1)
        wbt=w_frac(lowsum_pos,2)
        wct=w_frac(lowsum_pos,3)



function [dF]=solver_iso(w,F,k1,k2,k3,K1e,K2e,K3e,qo,Po,P)
% Solves the system of ODEs - Add the isomerisation rate constant
q=qo*(Po/P);%Delta=0 and Isothermal
%% Rate Expressions: - Isomerisation Assumed Irreversible
r1f=k1*F(1)/q;
r2f=k2*F(1)/q;
r3f=k3*F(2)/q;
r1r=(k1/K1e)*F(2)/q;
r2r=(k2/K2e)*F(3)/q;
r3r=(k3/K3e)*F(3)/q;
%% Solving the ODEs
    dF=zeros(3,1); % Each row for a differential equation
dF(1)=-(r1f+r2f)+(r1r+r2r);
dF(2)=-(r1r+r3f)+(r1f+r3r);
dF(3)=-(r2r+r3r)+(r2f+r3f);
end
```

Appendix E2: The Metathesis Model [Python]

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt


df = pd.read_excel("data.xlsx",sheet_name="simulation",header=0)
df_dist = pd.read_excel("Results_Summary_WO3_SiO2.xlsx",sheet_name="data_run1_s1",header=0)
target = np.array(df_dist['Total Flows [mol/s]'])

lump_target = np.sum(target[8::]) # Lumps all values in experimental data for >=C10
target[8] = lump_target # Lumps all values in experimental data for >=C10 into C10 position


def rate_const(n1,n2,ka,kb,kc,kd,ke):
    # Sets the value of the rate constant (k1,k4,k5,k8) based on the group
    # n1=k1/k8 - n2=M1
    # n1=k4/k5 - n2=M2
    if n1=='A':
        if n2=='A':
            k=ka[0]
        elif n2=='B':
            k=ka[1]
        elif n2=='C':
            k=ka[2]
        elif n2=='D':
            k=ka[3]
        else:
            k=ka[4]

    elif n1=='B':
        if n2=='A':
            k=kb[0]
        elif n2=='B':
            k=kb[1]
        elif n2=='C':
            k=kb[2]
        elif n2=='D':
            k=kb[3]
        else:
            k=kb[4]

    elif n1=='C':
        if n2=='A':
            k=kc[0]
        elif n2=='B':
            k=kc[1]
        elif n2=='C':
            k=kc[2]
        elif n2=='D':
            k=kc[3]
        else:
            k=kc[4]

    elif n1=='D':
        if n2=='A':
            k=kd[0]
```

```
        elif n2=='B':
            k=kd[1]
        elif n2=='C':
            k=kd[2]
        elif n2=='D':
            k=kd[3]
        else:
            k=kd[4]

    else:
        if n2=='A':
            k=ke[0]
        elif n2=='B':
            k=ke[1]
        elif n2=='C':
            k=ke[2]
        elif n2=='D':
            k=ke[3]
        else:
            k=ke[4]

    return k
# ------------------------------------------------------------------------

def metathesis_rates(conc,df,ka,kb,kc,kd,ke,T):
    # calculates rates for each component in a single reaction
    rates=np.zeros(15) # Return vector

    for i in range(0,df.shape[0]):

        k1=rate_const(df['k1'][i],df['m1'][i],ka,kb,kc,kd,ke)
        k8=rate_const(df['k8'][i],df['m1'][i],ka,kb,kc,kd,ke)
        k5=rate_const(df['k5'][i],df['m2'][i],ka,kb,kc,kd,ke)
        k4=rate_const(df['k4'][i],df['m2'][i],ka,kb,kc,kd,ke)

        # Select Equilibrium Constant Based on Temp [K]
        if T==693.15:
            Keq=df['Keq (693,15K)'][i]
        elif T==713.15:
            Keq=df['Keq (713,15K)'][i]
        else:
            Keq=df['Keq (733,15K)'][i]


        K = np.sqrt((Keq)/((k1/k4)*(k5/k8))) # Modified definition of ratios: k3/k2 = k7/k6

        # Special Terms
        k2_k23=(1/(K+1))
        k3_k23=(K/(K+1))

        k6_k67=k2_k23
        k7_k67=k3_k23



        # Hypothesis 3 Terms
#       kf = (k1 + k5)
#       kr = (kf/Keq)
```

```
#       kr = (k4 + k8) - changes depending on the hypothesis chosen
#       print(kf)

        # Component positions (in conc array)
        nr1=df['r1'][i]-1 # -1 as 0 indexed
        nr2=df['r2'][i]-1

        np1=df['p1'][i]-1
        np2=df['p2'][i]-1

#       # Rates - Hyp 3
#       Rr1 = kf*conc[nr1]*conc[nr2] - kr*conc[np1]*conc[np2]
#       Rr2 = Rr1
#
#       Rp1 = kr*conc[np1]*conc[np2] - kf*conc[nr1]*conc[nr2]
#       Rp2 = Rp1

        # Rates
        Rr1=k2_k23*(k1*conc[nr1] + k4*conc[np1]) - k1*conc[nr1]
        Rr2=k6_k67*(k5*conc[nr2] + k8*conc[np2]) - k5*conc[nr2]

        Rp1=k3_k23*(k1*conc[nr1] + k4*conc[np1]) - k4*conc[np1]
        Rp2=k7_k67*(k5*conc[nr2] + k8*conc[np2]) - k8*conc[np2]

        # Store the net rates
        rates[nr1]+=Rr1
        rates[nr2]+=Rr2
        rates[np1]+=Rp1
        rates[np2]+=Rp2

    return rates
# ------------------------------------------------------------------------------

def isomerisation_rates(conc,T):
    # T in Kelvin
    # conc @ 0 = 1-Hex
    # conc @ 1 = 2-Hex
    # conc @ 2 = 3-Hex

    rates_isom = np.zeros(3)

    # Equilibrium Constants - See Excel document: isomerisation_data.xlsx
    K1e = 8e-5*(T**2) - 0.1145*(T) + 42.472
    K2e = 4e-5*(T**2) - 0.0536*(T) + 19.987
    K3e = 1e-8*(T**2) - 5e-5*(T) + 0.204

    # Arrhenius Expressions - See Excel document: isomerisation_data.xlsx
    R=8.314 # J/mol/K
    k1 = 556.4626*np.exp(-62365.95/(R*T))
    k2 = 43105.25*np.exp(-90918.8/(R*T))
    k3 = 7697452.593*np.exp(-112762.78/(R*T))

    # Concentrations
    ca = conc[0] # 1-Hex
    cb = conc[1] # 2-Hex
    cc = conc[2] # 3-Hex

    # Rate Expressions
```

```
    r1f = k1*ca
    r1r = (k1/K1e)*cb

    r2f = k2*ca
    r2r = (k2/K2e)*cc

    r3f = k3*cb
    r3r = (k3/K3e)*cc

    # Store as species conversion rates
    rates_isom[0] = r1r + r2r - (r1f + r2f)
    rates_isom[1] = r1f + r3r - (r1r + r3f)
    rates_isom[2] = r2f + r3f - (r2r + r3r)

    return rates_isom

# ------------------------------------------------------------------------------

def dC(conc,w,ka,kb,kc,kd,ke,T,Qo):
    # Add on the isomerisation

    rates=metathesis_rates(conc,df,ka,kb,kc,kd,ke,T)
    rates=rates/Qo #


    rates_isom=isomerisation_rates(conc,T)
    rates_isom=rates_isom/Qo

    d = np.zeros(15)

    #Differential Equations
    d[0]=rates[0] + rates_isom[0]
    d[1]=rates[1] + rates_isom[1]
    d[2]=rates[2] + rates_isom[2]
    d[3]=rates[3]
    d[4]=rates[4]
    d[5]=rates[5]
    d[6]=rates[6]
    d[7]=rates[7]
    d[8]=rates[8]
    d[9]=rates[9]
    d[10]=rates[10]
    d[11]=rates[11]
    d[12]=rates[12]
    d[13]=rates[13]
    d[14]=rates[14]

    return d

#------------------------------------------------------------------------------

def l2_norm_calc(flows_out, target):
    # Lumps model predictions into product groups
    # Calculates the L2 norm of the (true - simulated) - Goal: Minimise L2 Norm
    comp_dist = np.zeros(15)

    comp_dist[0] = flows_out[3] #C2
    comp_dist[1] = flows_out[7] # C3
```

```
    comp_dist[2] = flows_out[9] + flows_out[11] # C4
    comp_dist[3] = flows_out[5] + flows_out[13] # C5
    comp_dist[4] = flows_out[0] + flows_out[1] + flows_out[2] # C6
    comp_dist[5] = flows_out[6] + flows_out[14] #C7
    comp_dist[6] = flows_out[10] + flows_out[12] #C8
    comp_dist[7] = flows_out[8] # C9
    comp_dist[8] = flows_out[4] #C10

    mask = comp_dist>0 # True wherever we have values above zero

    norm = np.sum(np.divide(np.sqrt(np.power((target[mask] - comp_dist[mask]),2)), comp_dist[mask])) # error vector
    print(norm)


    return norm, comp_dist
#------------------------------------------------------------------------------

# Solving ODE
from scipy.integrate import odeint
wf=0.001 # kg take from run 1 data
dw=0.000001
dvs = 25 # Number of decision variables
w=np.arange(0,wf,dw)
Cgo=10.28486413 # Take from run 1 data
Qo=5.02024e-6 # m^3/s
ic=[Cgo,0,0,0,0,0,0,0,0,0,0,0,0,0,0] # Initial Condition concentrations
T=693.15 #
# T = 693.15K or 713.15K or 733.15K


df_est = pd.read_excel("initial_constants.xlsx",sheet_name="s1",header=0)
params = np.array(df_est['constants'])

conc_out                                                                                    =
odeint(dC,ic,w,args=(params[0:5],params[5:10],params[10:15],params[15:20],params[20:25],T,Qo))
#[mol/m^3] - Contains matrix of values for each w-value
flows_out = conc_out*Qo #[mol/s]

norm, comp_dist = l2_norm_calc(flows_out[-1,:], target)
```

Appendix E3: Particle Swarm Optimization Implementation [Python]

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df = pd.read_excel("data.xlsx",sheet_name="simulation",header=0)
df_dist = pd.read_excel("Results_Summary_WO3_SiO2.xlsx",sheet_name="data_run1_s1",header=0)
target = np.array(df_dist['Total Flows [mol/s]'])

lump_target = np.sum(target[8::]) # Lumps all values in experimental data for >=C10
target[8] = lump_target # Lumps all values in experimental data for >=C10 into C10 position


def rate_const(n1,n2,x):
    # Sets the value of the rate constant (k1,k4,k5,k8) based on the group
    # n1=k1/k8 - n2=M1
    # n1=k4/k5 - n2=M2

    ka = x[0:5]
    kb = x[5:10]
    kc = x[10:15]
    kd = x[15:20]
    ke = x[20:25]

    if n1=='A':
        if n2=='A':
            k=ka[0]
        elif n2=='B':
            k=ka[1]
        elif n2=='C':
            k=ka[2]
        elif n2=='D':
            k=ka[3]
        else:
            k=ka[4]

    elif n1=='B':
        if n2=='A':
            k=kb[0]
        elif n2=='B':
            k=kb[1]
        elif n2=='C':
            k=kb[2]
        elif n2=='D':
            k=kb[3]
        else:
            k=kb[4]

    elif n1=='C':
        if n2=='A':
            k=kc[0]
        elif n2=='B':
            k=kc[1]
        elif n2=='C':
            k=kc[2]
        elif n2=='D':
            k=kc[3]
```

```
        else:
            k=kc[4]

    elif n1=='D':
        if n2=='A':
            k=kd[0]
        elif n2=='B':
            k=kd[1]
        elif n2=='C':
            k=kd[2]
        elif n2=='D':
            k=kd[3]
        else:
            k=kd[4]

    else:
        if n2=='A':
            k=ke[0]
        elif n2=='B':
            k=ke[1]
        elif n2=='C':
            k=ke[2]
        elif n2=='D':
            k=ke[3]
        else:
            k=ke[4]

    return k
# ------------------------------------------------------------------------------

def metathesis_rates(conc,df,x,T):
    # calculates rates for each component in a single reaction
    rates=np.zeros(15) # Return vector

    for i in range(0,df.shape[0]):

        k1=rate_const(df['k1'][i],df['m1'][i],x)
        k8=rate_const(df['k8'][i],df['m1'][i],x)
        k5=rate_const(df['k5'][i],df['m2'][i],x)
        k4=rate_const(df['k4'][i],df['m2'][i],x)

        # Select Equilibrium Constant Based on Temp [K]
        if T==693.15:
            Keq=df['Keq (693,15K)'][i]
        elif T==713.15:
            Keq=df['Keq (713,15K)'][i]
        else:
            Keq=df['Keq (733,15K)'][i]

        K = np.sqrt((Keq)/((k1/k4)*(k5/k8))) # Modified definition of ratios: k3/k2 = k7/k6


        # Hypothesis 2 Terms
        k2_k23=(1/(K+1))
        k3_k23=(K/(K+1))

        k6_k67=k2_k23
        k7_k67=k3_k23
```

```
        # Component positions (in conc array)
        nr1=df['r1'][i]-1 # -1 as 0 indexed
        nr2=df['r2'][i]-1

        np1=df['p1'][i]-1
        np2=df['p2'][i]-1

        # Rates
        Rr1=k2_k23*(k1*conc[nr1] + k4*conc[np1]) - k1*conc[nr1]
        Rr2=k6_k67*(k5*conc[nr2] + k8*conc[np2]) - k5*conc[nr2]

        Rp1=k3_k23*(k1*conc[nr1] + k4*conc[np1]) - k4*conc[np1]
        Rp2=k7_k67*(k5*conc[nr2] + k8*conc[np2]) - k8*conc[np2]

        # Store the net rates
        rates[nr1]+=Rr1
        rates[nr2]+=Rr2
        rates[np1]+=Rp1
        rates[np2]+=Rp2

    return rates
# ------------------------------------------------------------------------
def isomerisation_rates(conc,T):
    # T in Kelvin
    # conc @ 0 = 1-Hex
    # conc @ 1 = 2-Hex
    # conc @ 2 = 3-Hex

    rates_isom = np.zeros(3)

    # Equilibrium Constants - See Excel document: isomerisation_data.xlsx
    K1e = 8e-5*(T**2) - 0.1145*(T) + 42.472
    K2e = 4e-5*(T**2) - 0.0536*(T) + 19.987
    K3e = 1e-8*(T**2) - 5e-5*(T) + 0.204

    # Arrhenius Expressions - See Excel document: isomerisation_data.xlsx
    R=8.314 # J/mol/K
    k1 = 556.4626*np.exp(-62365.95/(R*T))
    k2 = 43105.25*np.exp(-90918.8/(R*T))
    k3 = 7697452.593*np.exp(-112762.78/(R*T))

    # Concentrations
    ca = conc[0] # 1-Hex
    cb = conc[1] # 2-Hex
    cc = conc[2] # 3-Hex

    # Rate Expressions
    r1f = k1*ca
    r1r = (k1/K1e)*cb

    r2f = k2*ca
    r2r = (k2/K2e)*cc

    r3f = k3*cb
    r3r = (k3/K3e)*cc

    # Store as species conversion rates
```

```
    rates_isom[0] = r1r + r2r - (r1f + r2f)
    rates_isom[1] = r1f + r3r - (r1r + r3f)
    rates_isom[2] = r2f + r3f - (r2r + r3r)

    return rates_isom

# ------------------------------------------------------------------------

def dC(conc,wm,x,T,Qo):
    # Add on the isomerisation
    rates=metathesis_rates(conc, df, x, T)
    rates=rates/Qo #


    rates_isom=isomerisation_rates(conc,T)
    rates_isom=rates_isom/Qo

    #Differential Equations
    d = np.zeros(15)

    #Differential Equations
    d[0]=rates[0] + rates_isom[0]
    d[1]=rates[1] + rates_isom[1]
    d[2]=rates[2] + rates_isom[2]
    d[3]=rates[3]
    d[4]=rates[4]
    d[5]=rates[5]
    d[6]=rates[6]
    d[7]=rates[7]
    d[8]=rates[8]
    d[9]=rates[9]
    d[10]=rates[10]
    d[11]=rates[11]
    d[12]=rates[12]
    d[13]=rates[13]
    d[14]=rates[14]

    return d

# ------------------------------------------------------------------------

def l2_norm_calc(flows_out, target):
    # Lumps model predictions into product groups
    # Calculates the L2 norm of the (true - simulated) - Goal: Minimise L2 Norm
    comp_dist = np.zeros(15)

    comp_dist[0] = flows_out[3] #C2
    comp_dist[1] = flows_out[7] # C3
    comp_dist[2] = flows_out[9] + flows_out[11] # C4
    comp_dist[3] = flows_out[5] + flows_out[13] # C5
    comp_dist[4] = flows_out[0] + flows_out[1] + flows_out[2] # C6
    comp_dist[5] = flows_out[6] + flows_out[14] #C7
    comp_dist[6] = flows_out[10] + flows_out[12] #C8
    comp_dist[7] = flows_out[8] # C9
    comp_dist[8] = flows_out[4] #C10

    mask = comp_dist>0 # True wherever we have values above zero
```

```
    norm = np.sum(np.divide(np.sqrt(np.power((target[mask] - comp_dist[mask]),2)), comp_dist[mask])) # ----------

    return norm, comp_dist
# ------------------------------------------------------------------------------

def search(target, vector):
    # Finds position of target in vector
    pos = 0
    for i in range(0,len(vector[0,:])): # ***
        if vector[0,i]==target:
            pos = i

    return pos

#-------------------------------------------------------------------------------

def generate_input(xo, n_params, N, c):
    # Creating a uniform initial position matrix
    x_mat = np.zeros( (n_params, N) )

    for i in range(0,n_params):
        min = np.floor(np.log10(xo[i])) - c # ***
        max = c + np.floor(np.log10(xo[i])) # ***

        x_mat[i,:] = np.reshape(np.append(np.logspace(min, max, (N-1)), xo[i]), (1, N ))

    return x_mat

#-------------------------------------------------------------------------------

def pso(N, c1, c2, w, num_iter, n_params, target, Qo, ic, w_cat, T):
    # Initial Parameters: (Position Vector - a estimates)
    df_est = pd.read_excel("initial_constants.xlsx",sheet_name="s1",header=0)

    xo = np.array(df_est['constants']) #np.reshape(np.logspace(-5,-8,n_params), (n_params,1)) # 25x1 - Read in from
Excel must be exponent x*10^y

#    print(xo)

    c = 1 # Controls the range of the search space from 0 to 1

    xo = generate_input(xo, n_params, N, c)

    # Random Components - Fixed Throughout
    r1 = np.random.rand(n_params,N) # Individual - 25xN particles
    r2 = np.random.rand(n_params,N) # Group - 25xN particles

    # Initial Velocity Vector: (Stochastic Parameter Change Amount)
    vo = 1e-13*np.random.rand(n_params,N)  # 25xN Vector  - one order below smallest

    iter_count = 1

    while iter_count < num_iter:
        if iter_count==1:

            x1 = xo
            v1 = vo
            # Calculating Objective: Minimise L1 Norm - Output Prediction Error Sum
```

```
        # Note that the objective functions will only have a single entry for each particle (regardless of number of params)
        x1_obj = []
        for i in range(0,N): # N particles

            flows_out = Qo * odeint(dC,ic,w_cat,args=(x1[:,i],T,Qo)) # mol/s

            l2_norm, comp_dist = l2_norm_calc(flows_out[-1,:], target) # L2 Norm - Based on mol/s flows

            x1_obj.append(l2_norm)

        x1_obj = np.reshape(np.array(x1_obj),(1,N)) # (1 row,N columns) - Always

        # Individual Best [1st iteration]
        pbest = x1 # Contains the best positions (params) for each particle

        pbest_obj = x1_obj # (1 row,N columns) - Always

        # Group Best [1st iteration]
        gbest_obj = np.min(pbest_obj) # (1 row,1 columns) - Always

        gbest = np.reshape(pbest[:,search(gbest_obj, pbest_obj)], (n_params,1)) # Search for best particle & update gbest (multiparam)


    else: # Iteration 2 to num_iter

        # Calculating Objective: Minimise L2 Norm - Output Prediction Error Sum
        x1_obj = []
        for i in range(0,N):

            flows_out = Qo * odeint(dC,ic,w_cat,args=(x1[:,i],T,Qo)) # mol/s

            l2_norm, comp_dist = l2_norm_calc(flows_out[-1,:], target) #np.sum(np.sqrt(np.power((y-y_pred), 2)))

            x1_obj.append(l2_norm)

        x1_obj = np.reshape(np.array(x1_obj), (1,N)) # Objective Function @ current iteration

        # pbest currently has the best params @ previous iteration

        # Updating pbest
        for i in range(0,N): # Over N Particles - across columns

            if x1_obj[0,i]<pbest_obj[0,i]: # Desire smallest L1 norm --- Not NaN conditions
                pbest[:,i] = x1[:,i] # 2 or more params here
                pbest_obj[0,i] = x1_obj[0,i]

            if pbest_obj[0,i]<gbest_obj: #--- Not NaN conditions
                gbest_obj = pbest_obj[0,i]
                gbest = np.reshape(pbest[:,search(gbest_obj, pbest_obj)], (n_params,1)) # multiple parameters [px1] in a particle


    # Velocity and Position calc at next iteration - For each particle:
    v2 = np.array(list())
#       print(v2)
    for i in range(0,n_params): # 2 or more parameters
```

```
        for j in range(0,N): # Over N Particles - across columns
            val = w[iter_count]*v1[i,j] + c1*r1[i,j]*(pbest[i,j] - x1[i,j]) + c2*r2[i,j]*(gbest[i,0] - x1[i,j])
            v2 = np.append(v2, val)

    v2 = np.reshape(v2, (n_params,N)) # currently stored as a row vector of length (n_params*N)

    x2 = x1 + v2 # Array sum

    #print(x2)

    iter_count+=1

    # Switch Data - as the next iteration became the current iteration

    v1 = v2
    x1 = x2

  return gbest, gbest_obj, flows_out, comp_dist


#-------------------------------------------------------------------------------
# Solving ODE
from scipy.integrate import odeint
wf = 0.001 # kg take from run 1 data
dw = wf/100
w_cat = np.arange(0,wf,dw)
Cgo = 10.28486413 # Take from run 1 data
Qo = 5.02024e-6 # m^3/s
ic=[Cgo,0,0,0,0,0,0,0,0,0,0,0,0,0,0] # Initial Condition concentrations
T = 693.15 #
# T = 693.15K or 713.15K or 733.15K

N = 50 # Number of particles (parameter sets)
c1 = 0.01 # Should not be larger than 1- Causes negative rate constants if too large#
c2 = 0.08 #
n_params = 25 # Numer of parameters
num_iter = 20 # Number of iterations
w = np.linspace(0.9,0.4,num_iter) # Inertia

params, l2_norm, flows_out, comp_dist = pso(N, c1, c2, w, num_iter,n_params,target, Qo, ic, w_cat, T) # Calling
PSO Algorithm
print(f'L2-Norm: {l2_norm}')
print(f'Parameters:\n {list(params)}')

# Write output ***
df_1 = pd.DataFrame(params)
df_1.to_excel("parameters_out.xlsx",sheet_name="results")

df_1 = pd.DataFrame(np.array(l2_norm).reshape(1,1))
df_1.to_excel("fitness_out.xlsx",sheet_name="results")

df_1 = pd.DataFrame(comp_dist)
df_1.to_excel("flows_out.xlsx",sheet_name="results")
```

Appendix E4: Genetic Algorithm Implementation [Python]

```
# Genetic Algorithm
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.integrate import odeint

df = pd.read_excel("data.xlsx",sheet_name="simulation",header=0)
df_dist = pd.read_excel("Results_Summary_WO3_SiO2.xlsx",sheet_name="data_run1_s1",header=0)
target = np.array(df_dist['Total Flows [mol/s]'])

def rate_const(n1,n2,x):
    # Sets the value of the rate constant (k1,k4,k5,k8) based on the group
    # n1=k1/k8 - n2=M1
    # n1=k4/k5 - n2=M2

    ka = x[0,0:5]
    kb = x[0,5:10]
    kc = x[0,10:15]
    kd = x[0,15:20]
    ke = x[0,20:25]

    if n1=='A':
        if n2=='A':
            k=ka[0]
        elif n2=='B':
            k=ka[1]
        elif n2=='C':
            k=ka[2]
        elif n2=='D':
            k=ka[3]
        else:
            k=ka[4]

    elif n1=='B':
        if n2=='A':
            k=kb[0]
        elif n2=='B':
            k=kb[1]
        elif n2=='C':
            k=kb[2]
        elif n2=='D':
            k=kb[3]
        else:
            k=kb[4]

    elif n1=='C':
        if n2=='A':
            k=kc[0]
        elif n2=='B':
            k=kc[1]
        elif n2=='C':
            k=kc[2]
        elif n2=='D':
            k=kc[3]
        else:
            k=kc[4]
```

```
    elif n1=='D':
        if n2=='A':
            k=kd[0]
        elif n2=='B':
            k=kd[1]
        elif n2=='C':
            k=kd[2]
        elif n2=='D':
            k=kd[3]
        else:
            k=kd[4]

    else:
        if n2=='A':
            k=ke[0]
        elif n2=='B':
            k=ke[1]
        elif n2=='C':
            k=ke[2]
        elif n2=='D':
            k=ke[3]
        else:
            k=ke[4]

    return k
# ----------------------------------------------------------------------------

def metathesis_rates(conc,df,x,T):
    # calculates rates for each component in a single reaction
    rates=np.zeros(15) # Return vector

    for i in range(0,df.shape[0]):

        k1=rate_const(df['k1'][i],df['m1'][i],x)
        k8=rate_const(df['k8'][i],df['m1'][i],x)
        k5=rate_const(df['k5'][i],df['m2'][i],x)
        k4=rate_const(df['k4'][i],df['m2'][i],x)

        # Select Equilibrium Constant Based on Temp [K]
        if T==693.15:
            Keq=df['Keq (693,15K)'][i]
        elif T==713.15:
            Keq=df['Keq (713,15K)'][i]
        else:
            Keq=df['Keq (733,15K)'][i]

#       # Hypothesis 3 Terms
#       kf = (k1 + k5)
#       kr = (kf/Keq)
##      kr = (k4 + k8)
##      print(kf)

        K = np.sqrt((Keq)/((k1/k4)*(k5/k8))) # Modified definition of ratios: k3/k2 = k7/k6

        # Hypothesis 2 Terms
        k2_k23=(1/(K+1))
        k3_k23=(K/(K+1))
```

```
        k6_k67=k2_k23
        k7_k67=k3_k23

        # Component positions (in conc array)
        nr1=df['r1'][i]-1 # -1 as 0 indexed
        nr2=df['r2'][i]-1

        np1=df['p1'][i]-1
        np2=df['p2'][i]-1

        # Rates
        Rr1=k2_k23*(k1*conc[nr1] + k4*conc[np1]) - k1*conc[nr1]
        Rr2=k6_k67*(k5*conc[nr2] + k8*conc[np2]) - k5*conc[nr2]

        Rp1=k3_k23*(k1*conc[nr1] + k4*conc[np1]) - k4*conc[np1]
        Rp2=k7_k67*(k5*conc[nr2] + k8*conc[np2]) - k8*conc[np2]

        # Store the net rates
        rates[nr1]+=Rr1
        rates[nr2]+=Rr2
        rates[np1]+=Rp1
        rates[np2]+=Rp2

    return rates
# ------------------------------------------------------------------------
def isomerisation_rates(conc,T):
    # T in Kelvin
    # conc @ 0 = 1-Hex
    # conc @ 1 = 2-Hex
    # conc @ 2 = 3-Hex

    rates_isom = np.zeros(3)

    # Equilibrium Constants - See Excel document: isomerisation_data.xlsx
    K1e = 8e-5*(T**2) - 0.1145*(T) + 42.472
    K2e = 4e-5*(T**2) - 0.0536*(T) + 19.987
    K3e = 1e-8*(T**2) - 5e-5*(T) + 0.204

    # Arrhenius Expressions - See Excel document: isomerisation_data.xlsx
    R=8.314 # J/mol/K
    k1 = 556.4626*np.exp(-62365.95/(R*T))
    k2 = 43105.25*np.exp(-90918.8/(R*T))
    k3 = 7697452.593*np.exp(-112762.78/(R*T))

    # Concentrations
    ca = conc[0] # 1-Hex
    cb = conc[1] # 2-Hex
    cc = conc[2] # 3-Hex

    # Rate Expressions
    r1f = k1*ca
    r1r = (k1/K1e)*cb

    r2f = k2*ca
    r2r = (k2/K2e)*cc

    r3f = k3*cb
```

```
    r3r = (k3/K3e)*cc

    # Store as species conversion rates
    rates_isom[0] = r1r + r2r - (r1f + r2f)
    rates_isom[1] = r1f + r3r - (r1r + r3f)
    rates_isom[2] = r2f + r3f - (r2r + r3r)

    return rates_isom

# ------------------------------------------------------------------------

def dC(conc,wm,x,T,Qo):
    # Add on the isomerisation
    rates=metathesis_rates(conc, df, x, T)
    rates=rates/Qo #

    rates_isom=isomerisation_rates(conc,T)
    rates_isom=rates_isom/Qo

    #Differential Equations
    d = np.zeros(15)

    #Differential Equations
    d[0]=rates[0] + rates_isom[0]
    d[1]=rates[1] + rates_isom[1]
    d[2]=rates[2] + rates_isom[2]
    d[3]=rates[3]
    d[4]=rates[4]
    d[5]=rates[5]
    d[6]=rates[6]
    d[7]=rates[7]
    d[8]=rates[8]
    d[9]=rates[9]
    d[10]=rates[10]
    d[11]=rates[11]
    d[12]=rates[12]
    d[13]=rates[13]
    d[14]=rates[14]

    return d
# ------------------------------------------------------------------------

def l2_norm_calc(flows_out, target, best_norm):
    # Lumps model predictions into product groups
    # Calculates the L2 norm of the (true - simulated) - Goal: Minimise L2 Norm
    comp_dist = np.zeros(15)

    comp_dist[0] = flows_out[3] #C2
    comp_dist[1] = flows_out[7] # C3
    comp_dist[2] = flows_out[9] + flows_out[11] # C4
    comp_dist[3] = flows_out[5] + flows_out[13] # C5
    comp_dist[4] = flows_out[0] + flows_out[1] + flows_out[2] # C6
    comp_dist[5] = flows_out[6] + flows_out[14] #C7
    comp_dist[6] = flows_out[10] + flows_out[12] #C8
    comp_dist[7] = flows_out[8] # C9
    comp_dist[8] = flows_out[4] #C10

    mask = comp_dist>0 # True wherever we have values above zero
```

```
    norm = np.sum(np.divide(np.sqrt(np.power((target[mask] - comp_dist[mask]),2)), comp_dist[mask])) # error
vector

    if best_norm==0:
        best_norm = norm

    return norm, comp_dist, best_norm

# ------------------------------------------------------------------------
def random_population(Np, dvs): #***
    # Np - Number of sets in the population - rows
    # dvs - NUmber of DVs in the fitness funtion - columns

    df_est = pd.read_excel("initial_constants.xlsx",sheet_name="s1",header=0)
    x = np.reshape(np.array(df_est['constants']),(1,dvs)) # Unshaped - 25 values - Seed for random population

    # Need to adjust this for decimal values -  could just read in samples from excel
    P = 10*np.random.rand(Np, dvs) # Random population - range 1 to 10(excl)

    for i in range(0,dvs):

        P[:,i] = x[0,i]*P[:,i] # Each row contains a new set of DVs

    return P
# ------------------------------------------------------------------------
def fitness(P, f, Np, dvs, Qo, ic, w_cat, T, best_norm): #***
    # Qo, ic, w_cat, T
    for i in range(0,Np): # For each DV in specific set
        flows_out = Qo * odeint(dC,ic,w_cat,args=( np.reshape(P[i,:],(1,dvs)) ,T,Qo))
        l2_norm, comp_dist, best_norm = l2_norm_calc(flows_out[-1,:], target, best_norm)
        f[i,0] = l2_norm

    return f, flows_out, best_norm # ***
# ------------------------------------------------------------------------
def tournament(P, k, Np, fp, dvs):
     #np.arange(0,Np) #
    pool = np.zeros((Np, dvs))
    pool_fitness = np.zeros((Np, 1))
    count = 0 # Counts the entry position in mating pool

    for i in range(0,2): # 2 Tournaments
        T = P
        n = Np

        for i in range(0, int(Np/2)): # Tests two sets in one comparison so half the number of executions required

            mask = np.ones(n, dtype=bool) # Boolean mask
            pos = np.random.randint(0,n) # Position to be taken for competition - 0 to (Np-1)
            T1 = T[pos, :] # Participant A
            f1 = fp[pos,0] # Participant A Fitness
            mask[pos] = False
            T = T[mask] # Removing the set A from the population  - P still contains all sets
            n = n-1

            mask = np.ones(n, dtype=bool) # Boolean mask
            pos = np.random.randint(0,n) # Position to be taken for competition - 0 to (Np-1)
            T2 = T[pos, :] # Participant B
```

```
        f2 = fp[pos,0] # Participant B Fitness
        mask[pos] = False
        T = T[mask] # Removing the set B from the population  - P still contains all sets
        n = n-1

        if f1<f2: # This will change depending on objective (minimisation here)
            pool[count, :] = T1
            pool_fitness[count, :] = f1
            count = count + 1

        else:
            pool[count, :] = T2
            pool_fitness[count, :] = f2
            count = count + 1

    return pool, pool_fitness
#-------------------------------------------------------------------------------
def bounder(vec, lb, ub, dvs): #***
    # Bounds the generated values to lie in the domain
    # vec- column vector (array) with DV values along column
    # dvs - Number of dvs in the system
    for i in range(0,dvs):

        if vec[0,i]>ub:
            vec[0,i] = ub

        elif vec[0,i]<lb:
            vec[0,i] = lb

    return vec
#-------------------------------------------------------------------------------
def crossover(Np, P, pc, nc, dvs, lb, ub):
    offspring = np.zeros((Np, dvs))
    #offspring_fitness = np.zeros((Np, 1))

    T = P
    n = Np
    count = 0 # Counts the entry position in offspring array


    for i in range(0, int(Np/2)): # Tests two sets in one comparison so half the number of executions required
        r = np.random.rand() # Used to determine if crossover will occur

        # Choosing parents at random:
        mask = np.ones(n, dtype=bool) # Boolean mask
        pos = np.random.randint(0,n) # Position to be taken for competition - 0 to (Np-1)
        T1 = T[pos, :] # Parent A
        mask[pos] = False
        T = T[mask] # Removing the Parent A from the population  - P still contains all sets
        n = n-1

        mask = np.ones(n, dtype=bool) # Boolean mask
        pos = np.random.randint(0,n) # Position to be taken for competition - 0 to (Np-1)
        T2 = T[pos, :] # Parent B
        mask[pos] = False
        T = T[mask] # Removing the Parent B from the population  - P still contains all sets
        n = n-1
```

```
        if r<pc:
            # Apply SBX Formula:
            u = np.random.rand(1, dvs) # 1 x num_dvs array or random numbers
            beta = np.zeros((1,dvs))

            for i in range(0,dvs):
                if u[0,i]>0.5:
                    beta[0,i] = (1/(2*(1-u[0,i])))**(1/(nc+1))

                else:
                    beta[0,i] = (2*u[0,i])**(1/(nc+1))
            # New Offsprings - Vector Algebra
            O1 = 0.5*( (1+beta)*T1 + (1-beta)*T2 ) # 1xdvs
            O2 = 0.5*( (1-beta)*T1 + (1+beta)*T2 )

            # Bounding
            O1 = bounder(O1, lb, ub, dvs)
            O2 = bounder(O2, lb, ub, dvs)

            # Saving Offsprings in Output Vector
            offspring[count, :] = O1
            count = count + 1

            offspring[count, :] = O2
            count = count + 1

        else:
            # Copy parent solutions - Could save fitness as well and only calculate where offspring fitness is unchanged(0)
            offspring[count, :] = T1
            count = count + 1

            offspring[count, :] = T2
            count = count + 1

    return offspring
#-------------------------------------------------------------------------------
def mutation(Np, O, pm, nm, dvs, lb, ub):

    for i in range(0,Np):
        # Generate random number
        u = np.random.rand()

        if u<pm:
            r = np.random.rand(1,dvs)

            delta = np.zeros((1,dvs))

            for j in range(0,dvs): # DVs along columns

                if r[0,j]>=0.5:
                    delta[0,j] = 1 - ((2*(1-r[0,j]))**(1/(nm+1)))

                else:
                    delta[0,j] = ((2*r[0,j])**(1/(nm+1))) - 1

            # Offspring Mutation - Vector Algebra
            O1 = O[i,:] + (ub - lb)*delta # 1xdvs
```

```
            # Bounding
            O[i,:] = bounder(O1, lb, ub, dvs)

    return O
#-------------------------------------------------------------------------------
def survival(P, fp, O, fo, Np, dvs):
    new_population = np.concatenate((P, O), axis = 0)
    new_fitness = np.concatenate((fp, fo), axis = 0)

    pop_fit = np.concatenate((new_population,new_fitness), axis=1)

    pop_fit = pop_fit[pop_fit[:, dvs].argsort()] # Zero indexed so dvs = last column position
    # argsort sorts from smallest to biggest - since we are minimising the fitness, take top Np values
    P = pop_fit[0:Np, 0:dvs]
    f = np.reshape(pop_fit[0:Np, dvs], (Np,1)) # reshaped as the slicing makes it an ndarray

    return P, f
#-------------------------------------------------------------------------------
def genetic_alg(dvs, lb, ub, Np, iter_count, pc, pm, nc, nm, k, Qo, ic, w_cat, T):
    fp = np.zeros((Np, 1)) # Initialised population fitness array (Npx1)
    fo = np.zeros((Np, 1)) # Initialised offspring fitness array  (Npx1)

    best_norm = 0

    P = random_population(Np, dvs)

    fp, final_flows, best_norm = fitness(P, fp, Np, dvs, Qo, ic, w_cat, T, best_norm)

    for i in range(0,iter_count):
        # Tournament Selection
        P,fp = tournament(P, k, Np, fp, dvs) # P will now hold the mating pool

        # SBX Crossover
        O = crossover(Np, P, pc, nc, dvs, lb, ub)

        # Polynomial Mutation
        O = mutation(Np, O, pm, nm, dvs, lb, ub)

        # Offspring Fitness
        fo, final_flows, best_norm = fitness(O, fo, Np, dvs, Qo, ic, w_cat, T, best_norm)

        P, fp = survival(P, fp, O, fo, Np, dvs) # Updates fp

    return P[0,:], fp[0,0], final_flows

# Decision variable Bounds *** - Use PSO to find seed values and genetic to explore range
lb = 1e-9 #***
ub = 1e-3

# Operating Parameters
Np = 6 # Number of solutions in population - Even Number otherwise causes issues
iter_count = 5 # Number of iterations
dvs = 25 # Number of decision variables in fitness function

# Fitness f(x) - Find its minimum

# Probabilities
pc = 0.9 # Crossover
```

```
pm = 0.3 # Mutation

# Distribution Indices
nc = 2.5 #***
nm = 5

# Tournament
k = 2 # Number of solutions participating in tournament

# Other Parameters ----------
wf = 0.001 # kg take from run 1 data ----------
dw = wf/100 # ----------
w_cat = np.arange(0,wf,dw)
Cgo = 10.28486413 # Take from run 1 data ----------
Qo = 5.02024e-6 # m^3/s ----------
ic=[Cgo,0,0,0,0,0,0,0,0,0,0,0,0,0,0] # Initial Condition concentrations
T = 693.15
# T = 693.15K or 713.15K or 733.15K

# Call Genetic Algorithm
P_final, f_final, flows_final = genetic_alg(dvs, lb, ub, Np, iter_count, pc, pm, nc, nm, k, Qo, ic, w_cat, T)

flows_final = flows_final[-1,:] # Now contains the final flowrates of each species - Need for profiles

# Grouping Flows ***
comp_dist = np.zeros(15)

comp_dist[0] = flows_final[3] #C2
comp_dist[1] = flows_final[7] # C3
comp_dist[2] = flows_final[9] + flows_final[11] # C4
comp_dist[3] = flows_final[5] + flows_final[13] # C5
comp_dist[4] = flows_final[0] + flows_final[1] + flows_final[2] # C6
comp_dist[5] = flows_final[6] + flows_final[14] #C7
comp_dist[6] = flows_final[10] + flows_final[12] #C8
comp_dist[7] = flows_final[8] # C9
comp_dist[8] = flows_final[4] #C10

# Write output ***
df_1 = pd.DataFrame(P_final)
df_1.to_excel("parameters_out.xlsx",sheet_name="results")

df_1 = pd.DataFrame(np.array(f_final).reshape(1,1)) #*****
df_1.to_excel("fitness_out.xlsx",sheet_name="results")

df_1 = pd.DataFrame(comp_dist)
df_1.to_excel("flows_out.xlsx",sheet_name="results")
```

Appendix E5: Kinetic Model Implementation [Python]

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df = pd.read_excel("data.xlsx",sheet_name="simulation",header=0)
df_dist = pd.read_excel("Results_Summary_WO3_SiO2.xlsx",sheet_name="data_run1_s1",header=0)
target = np.array(df_dist['Total Flows [mol/s]'])

lump_target = np.sum(target[8::]) # Lumps all values in experimental data for >=C10
target[8] = lump_target # Lumps all values in experimental data for >=C10 into C10 position


def rate_const(n1,n2,E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T):
    # Sets the value of the rate constant (k1,k4,k5,k8) based on the group
    # n1=k1/k8 - n2=M1
    # n1=k4/k5 - n2=M2

    if n1=='A':
        if n2=='A':
            k=A1[0]*np.exp(-E1[0]/((8.3145)*(T)))
        elif n2=='B':
            k=A1[1]*np.exp(-E1[1]/((8.3145)*(T)))
        elif n2=='C':
            k=A1[2]*np.exp(-E1[2]/((8.3145)*(T)))
        elif n2=='D':
            k=A1[3]*np.exp(-E1[3]/((8.3145)*(T)))
        else:
            k=A1[4]*np.exp(-E1[4]/((8.3145)*(T)))

    elif n1=='B':
        if n2=='A':
            k=A2[0]*np.exp(-E2[0]/((8.3145)*(T)))
        elif n2=='B':
            k=A2[1]*np.exp(-E2[1]/((8.3145)*(T)))
        elif n2=='C':
            k=A2[2]*np.exp(-E2[2]/((8.3145)*(T)))
        elif n2=='D':
            k=A2[3]*np.exp(-E2[3]/((8.3145)*(T)))
        else:
            k=A2[4]*np.exp(-E2[4]/((8.3145)*(T)))

    elif n1=='C':
        if n2=='A':
            k=A3[0]*np.exp(-E3[0]/((8.3145)*(T)))
        elif n2=='B':
            k=A3[1]*np.exp(-E3[1]/((8.3145)*(T)))
        elif n2=='C':
            k=A3[2]*np.exp(-E3[2]/((8.3145)*(T)))
        elif n2=='D':
            k=A3[3]*np.exp(-E3[3]/((8.3145)*(T)))
        else:
            k=A3[4]*np.exp(-E3[4]/((8.3145)*(T)))

    elif n1=='D':
```

```
        if n2=='A':
            k=A4[0]*np.exp(-E4[0]/((8.3145)*(T)))
        elif n2=='B':
            k=A4[1]*np.exp(-E4[1]/((8.3145)*(T)))
        elif n2=='C':
            k=A4[2]*np.exp(-E4[2]/((8.3145)*(T)))
        elif n2=='D':
            k=A4[3]*np.exp(-E4[3]/((8.3145)*(T)))
        else:
            k=A4[4]*np.exp(-E4[4]/((8.3145)*(T)))

    else:
        if n2=='A':
            k=A5[0]*np.exp(-E5[0]/((8.3145)*(T)))
        elif n2=='B':
            k=A5[1]*np.exp(-E5[1]/((8.3145)*(T)))
        elif n2=='C':
            k=A5[2]*np.exp(-E5[2]/((8.3145)*(T)))
        elif n2=='D':
            k=A5[3]*np.exp(-E5[3]/((8.3145)*(T)))
        else:
            k=A5[4]*np.exp(-E5[4]/((8.3145)*(T)))

    return k
# ----------------------------------------------------------------------------

def metathesis_rates(conc,df,E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T):
    # calculates rates for each component in a single reaction
    rates=np.zeros(15) # Return vector

    for i in range(0,df.shape[0]):

        k1=rate_const(df['k1'][i],df['m1'][i],E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T)
        k8=rate_const(df['k8'][i],df['m1'][i],E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T)
        k5=rate_const(df['k5'][i],df['m2'][i],E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T)
        k4=rate_const(df['k4'][i],df['m2'][i],E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T)

        # Select Equilibrium Constant Based on Temp [K]
        if T==693.15:
            Keq=df['Keq (693,15K)'][i]
        elif T==713.15:
            Keq=df['Keq (713,15K)'][i]
        else:
            Keq=df['Keq (733,15K)'][i]


        K = np.sqrt((Keq)/((k1/k4)*(k5/k8))) # Modified definition of ratios: k3/k2 = k7/k6

        # Special Terms
        k2_k23=(1/(K+1))
        k3_k23=(K/(K+1))

        k6_k67=k2_k23
        k7_k67=k3_k23

        # Component positions (in conc array)
```

```
        nr1=df['r1'][i]-1 # -1 as 0 indexed
        nr2=df['r2'][i]-1

        np1=df['p1'][i]-1
        np2=df['p2'][i]-1

        # Rates
        Rr1=k2_k23*(k1*conc[nr1] + k4*conc[np1]) - k1*conc[nr1]
        Rr2=k6_k67*(k5*conc[nr2] + k8*conc[np2]) - k5*conc[nr2]

        Rp1=k3_k23*(k1*conc[nr1] + k4*conc[np1]) - k4*conc[np1]
        Rp2=k7_k67*(k5*conc[nr2] + k8*conc[np2]) - k8*conc[np2]

        # Store the net rates
        rates[nr1]+=Rr1
        rates[nr2]+=Rr2
        rates[np1]+=Rp1
        rates[np2]+=Rp2

    return rates
# ------------------------------------------------------------------------
def isomerisation_rates(conc,T):
    # T in Kelvin
    # conc @ 0 = 1-Hex
    # conc @ 1 = 2-Hex
    # conc @ 2 = 3-Hex

    rates_isom = np.zeros(3)

    # Equilibrium Constants - See Excel document: isomerisation_data.xlsx
    K1e = 8e-5*(T**2) - 0.1145*(T) + 42.472
    K2e = 4e-5*(T**2) - 0.0536*(T) + 19.987
    K3e = 1e-8*(T**2) - 5e-5*(T) + 0.204

    # Arrhenius Expressions - See Excel document: isomerisation_data.xlsx
    R=8.314 # J/mol/K
    k1 = 556.4626*np.exp(-62365.95/(R*T))
    k2 = 43105.25*np.exp(-90918.8/(R*T))
    k3 = 7697452.593*np.exp(-112762.78/(R*T))

    # Concentrations
    ca = conc[0] # 1-Hex
    cb = conc[1] # 2-Hex
    cc = conc[2] # 3-Hex

    # Rate Expressions
    r1f = k1*ca
    r1r = (k1/K1e)*cb

    r2f = k2*ca
    r2r = (k2/K2e)*cc

    r3f = k3*cb
    r3r = (k3/K3e)*cc

    # Store as species conversion rates
```

```
    rates_isom[0] = r1r + r2r - (r1f + r2f)
    rates_isom[1] = r1f + r3r - (r1r + r3f)
    rates_isom[2] = r2f + r3f - (r2r + r3r)

    return rates_isom

# ------------------------------------------------------------------------------

def dC(conc,w,E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T,Qo):
    # Add on the isomerisation

    rates=metathesis_rates(conc,df,E1,E2,E3,E4,E5,A1,A2,A3,A4,A5,T)
    rates=rates/Qo #


    rates_isom=isomerisation_rates(conc,T)
    rates_isom=rates_isom/Qo

    d = np.zeros(15)

    #Differential Equations
    d[0]=rates[0] + rates_isom[0]
    d[1]=rates[1] + rates_isom[1]
    d[2]=rates[2] + rates_isom[2]
    d[3]=rates[3]
    d[4]=rates[4]
    d[5]=rates[5]
    d[6]=rates[6]
    d[7]=rates[7]
    d[8]=rates[8]
    d[9]=rates[9]
    d[10]=rates[10]
    d[11]=rates[11]
    d[12]=rates[12]
    d[13]=rates[13]
    d[14]=rates[14]

    return d

#------------------------------------------------------------------------------
def l2_norm_calc(flows_out, target):
    # Lumps model predictions into product groups
    # Calculates the L2 norm of the (true - simulated) - Goal: Minimise L2 Norm
    comp_dist = np.zeros(15)

    comp_dist[0] = flows_out[3] #C2
    comp_dist[1] = flows_out[7] # C3
    comp_dist[2] = flows_out[9] + flows_out[11] # C4
    comp_dist[3] = flows_out[5] + flows_out[13] # C5
    comp_dist[4] = flows_out[0] + flows_out[1] + flows_out[2] # C6
    comp_dist[5] = flows_out[6] + flows_out[14] #C7
    comp_dist[6] = flows_out[10] + flows_out[12] #C8
    comp_dist[7] = flows_out[8] # C9
    comp_dist[8] = flows_out[4] #C10

    mask = comp_dist>0 # True wherever we have values above zero
```

```
    norm = np.sum(np.divide(np.sqrt(np.power((target[mask] - comp_dist[mask]),2)), comp_dist[mask])) # error vector
    print(norm)

    return norm, comp_dist

# Solving ODE
from scipy.integrate import odeint
wf=0.001 # kg take from run 1 data
dw=0.000001
dvs = 25 # Number of decision variables
w=np.arange(0,wf,dw)
Cgo=10.28486413 # Take from run 1 data
Qo=5.02024e-6 # m^3/s
ic=[Cgo,0,0,0,0,0,0,0,0,0,0,0,0,0,0] # Initial Condition concentrations
T=693.15
# T = 693.15K or 713.15K or 733.15K

df_est = pd.read_excel("initial_constants_model.xlsx",sheet_name="s1",header=0)
params_1 = np.array(df_est['Ea'])
params_2 = np.array(df_est['A'])


conc_out = odeint(dC,ic,w,args=(params_1[0:5],
                                params_1[5:10],
                                params_1[10:15],
                                params_1[15:20],
                                params_1[20:25],
                                params_2[0:5],
                                params_2[5:10],
                                params_2[10:15],
                                params_2[15:20],
                                params_2[20:25],
                                T,
                                Qo)) #[mol/m^3] - Contains matrix of values for each w-value
flows_out = conc_out*Qo #[mol/s]


norm, comp_dist = l2_norm_calc(flows_out[-1,:], target)


df_1 = pd.DataFrame(comp_dist)
df_1.to_excel("outputs_model.xlsx",sheet_name="results_model")
```