University of Alberta

Lumped kinetic modelling and multivariate data analysis of propylene conversion over H-ZSM-5

by

Jinjun Nie

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Chemical Engineering

Department of Chemical and Materials Engineering

Spring 2013
Edmonton, Alberta

Abstract

The process of catalytic conversion of propylene over ZSM-5 zeolite catalyst in a plug-flow micro-reactor with continuous feed flow is studied under near atmospheric pressure and in the temperature range of 330 ℃ to 450 ℃ with gas hourly velocity from 110/hour to 300/hour. The reaction product is condensed and characterized by gas-phase GC, liquid-phase GC, offline FTIR, online FTIR, NIR and UV-Vis spectroscopy. The effects of space velocity and reaction temperature on the carbon number distribution, the yield of C_1-C_2 gases, C_4 hydrocarbons, C_{5+} hydrocarbons and yield of aromatics are evaluated. The chemical species are lumped into C_1-C_2 light gas, propylene, propane, C_4 hydrocarbons, C_{5+} hydrocarbons and aromatics and a lumped kinetic model is built with the kinetic parameters estimated by the experimental data. The prediction of concentrations of the chemical lumps by the built model shows a good agreement with the experimental data from extra experiments, which further validate the practice of the model prediction. Multivariate data analysis, which includes Principal Components Analysis and Hierarchical Clustering Analysis, is employed for analysis of the experimental data and also the offline and online spectra data for further understanding the process. The mapping of the spectra in relation to the reaction conditions is built and validated by the experimental tests. Algorithms of outlier detection and process control of the reaction process are proposed, as well as the application of this mapping with multivariate data analysis is proposed for future work.

Acknowledgement

I would like to take this chance to thank Dr. Arno de Klerk and Dr. Vinay Prasad for their responsible and supportive supervision on my project work. It is their nice supervision that makes this thesis possible. Their valuable views in academic and meritorious suggestions on my lab work and data analysis help me a lot in completion of this study. Also I learn a lot from them both in thinking and writing in a scientific way and in time management which will be critically important in my future life.

Thanks also to Dr. Shaofeng Yang, who helped me a lot in my lab work and gave me a lot support in project management and equipment training. I also sincerely appreciate the Helmholtz-Alberta Initiative for its funding on my work. I would also appreciate my friends and my colleagues, it is their smile and their support that enriches these two and half years and makes my research life happier and colorful.

The sincerest thanks to my family, to my dear father and my dear mother, and to my lovely sister, who never failed to support me as they are always there for me. Nothing can stop their love to me, never the time, and never the several thousand miles across the Pacific Ocean. I am so grateful to have them in my life. I will be indebted, forever, from the bottom of my heart.

Last, I would like to give my thanks to all who helped me and who were present in my life giving me important lessons to lead or strive for a better life, and best wishes to all who are on their way of pursuit of happiness as now I am.

Table of Contents

List of Figures

List of Tables

Nomenclatures

k_{ij}	Reaction rate constant
$\overline{MW_i}$	Average molecular weight of the lump *I*
v_{ij}	Stoichiometric coefficient of reaction from Lump *I* to lump *J*
m_{acc}	Mass accumulation
m_{in}	Mass flow in
m_{out}	Mass flow out
m_{gen}	Mass generated
m_{con}	Mass consumption
C_i	Concentration of lump I
u	Linear velocity of reaction flow
dx	Length difference
dt	Time difference
P	Pressure of gas flow
R	Gas constant
n_t	Mole flow rate of the gas flow
T	Temperature in K
A	Cross area of the reactor
r_i	Reaction rate of lump I
V_t	Gas volume flow rate
A_0	Reaction rate frequency
E_a	Apparent activation energy

C_{ical} Concentration of lumps calculated by the model

C_{iexp} Concentration of lumps measured by analysis

f_{min} The objective value to minimize

Abbreviations

GC	Gas chromatography
MS	Mass spectrometry
FID	Flame ionization detector
TCD	Thermal conductor detector
IR	Infrared spectra
FTIR	Fourier transferred Infrared spectra
NIR	Near-infrared spectra
UV-Vis	Ultraviolet-visible spectra
HCA	Hierarchical Clustering Analysis
PCA	Principal Components Analysis
TGA	Thermal Gravimetric Analysis
WHSV	Weigh Hourly Space Velocity
ZSM-5	Zeolites Socony Mobil – 5
ODE	Ordinary Differential Equation

Chapter 1

Introduction and Literature Review

1.1 Introduction

1.1.1 Background

Energy is one of the most important components for the maintenance and development of industry and society worldwide. The balance of energy supply and consumption plays a critical role in supporting the healthy and sustainable operation of the whole society. Currently, oil, coal and natural gas are the three main global energy sources, which account for 33.2%, 27.0% and 21.1% of the overall energy demand respectively (IEA, 2010). The sources of oil are further classified into conventional crude oil which is normally defined as light and medium crude oil, non-conventional crude oil such as oil sands and oil shale, and natural gas liquids which is the liquid component of natural gas (Bentley, et al., 2009). The challenge facing the oil industry is how to increase production when existing conventional crude oil fields are experiencing depletion rate of 5% or more while annual demand for oil is expected to increase at rates exceeding 1% (Hedges et al., 2011). Thus growth in exploration and production of available unconventional heavy oil, which is abundant in certain areas and easy to mine technically, gets more focus. A good example is the mining of oil sands in Alberta, Canada.

As the oil industry is being forced to look for non-conventional sources, notably Canada's oil sands maintain the balance of mining and consumption of

oil (De Castro et al., 2009). Pyrolysis of these heavy oils from the oil sands becomes the main resource to produce light product in the range of light gas oil and distillate. Meantime, large amount of light gas in the pyrolysis process are also produced which is less valuable than the liquid fuel product.

1.1.2 Project introduction

Pyrolysis gas is the gas from the pyrolysis process which is a process of high-temperature thermo-chemical decomposition, commonly under pressure and operating temperature over 400℃, of certain organic materials. These organic materials extend to bitumen (Murugan et al., 2012) with the purpose of converting heavy molecules into more valuable lighter products. The pyrolysis gas that is co-produced is usually rich in hydrogen, light olefins, paraffins and carbon monoxide, the composition depending on the pyrolysis temperature. This gas mixture can be converted to light oil product or other valuable chemicals under certain thermal conditions with the help of a catalyst.

Our project focuses on the conversion of pyrolysis gas to a light oil product with the help of a ZSM-5 zeolite catalyst. Although our interest is in bitumen, the source of pyrolysis gas can be from the pyrolysis of other materials too, such as biomass. The present study investigates a simplified system and only pure propylene is used as feedstock in a continuous fix-bed flow reactor to study the oligomerization process and the side reactions that can happen. The results from this work, including thermal analysis, kinetic study, and spectroscopy analysis, offers important guidance in further work of conversion of pyrolysis gases from thermal cracking of bitumen. The specific objective of the work is to develop a reaction model to describe the reaction network so that it is linked to spectroscopic information that can in principle be obtained online. This will enable future studies to use on-line spectroscopy to track changes in the reaction network when employing real pyrolysis gas as feed. Real pyrolysis gas will likely accelerate catalyst deactivation, making real time online analysis the preferred method of following the reaction.

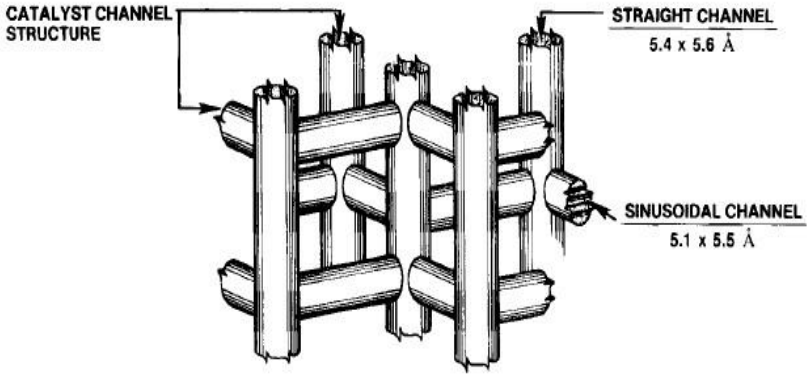
1.2 Literature review

1.2.1 Chemistry of light gas conversion

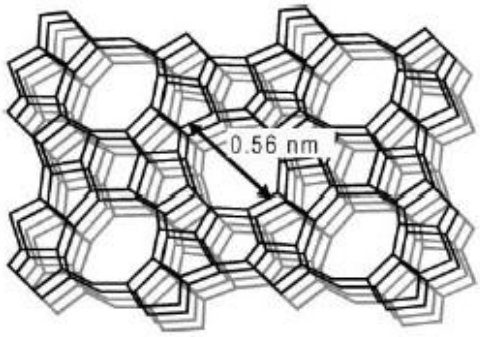
One of the most important uses of acid-type catalysts in oil refining is catalyzing the oligomerization of light olefins to higher molecular weight liquid products (Oblad et al., 1958). Earlier, in 1935, this chemistry was commercialized to produce gasoline-range iso-olefins from C3 and C4 olefins, using phosphoric acid impregnated on silica clay (Ipatieff and Egloff, 1934; Tabak et al., 1983). Later, in the period of 1960s to 1980s, effort was put into developing technology for olefin conversion over zeolite catalysts by Mobil Research and Development Corporation, which resulted in the Mobil Olefin to Gasoline and Distillate (MOGD) process. Using a ZSM-5 zeolites catalyst, this process converts light olefins to higher molecular weight products in the range of gasoline and diesel fuel.

This process chemistry is based on a unique synthetic crystalline material ZSM-5 (Kokotailo et al., 1978; Tabak et al., 1983). The catalytic reactions happen on the protonic acid sites within the catalyst, and it is the shape and diameter of the pores that control the shape and size of the product molecules (Tabak et al., 1983). It is well known that ZSM-5 catalysts have high selectivity on a more linear oligomerization product for the reactions due to its structure, which effectively constrains the molecular size and structure of the transition state in the reactions.

Figure 1-1 shows the structure of ZSM-5 zeolites catalysts (Kokotailo et al., 1978; Tabak et al., 1983; Hollander et al., 2002). Detailed information about ZSM-5 is described by Kokotailo et al.



(a)



(b)

Figure 1-1. Structure of ZSM-5 zeolite catalyst (a) The channel structure of ZSM-5 zeolites (Kokotailo et al., 1978; Tabak et al., 1983); (b) The sketch structure of ZSM-5 zeolites (Hollander et al., 2002; Baerlocher et al., 2000)

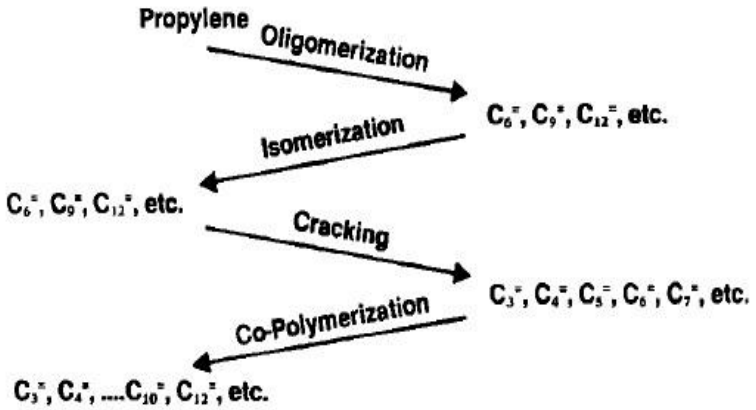


Figure 1-2. Reaction path for propylene conversion over ZSM-5 zeolites (Tabak et al., 1983)

One view as indicated by Tabak et al. in their work is that the reaction of light olefins over ZSM-5 catalyst was found to proceed by reaction to discrete oligomers, followed by cracking, copolymerization, self-polymerization and aromatization. The shape of the product molecules is governed primarily by the structure of the zeolite catalyst. The type and amount of chemical species in the product are affected by temperature, residence time and pressure. Certain reaction equilibrium constraints become important and limit the carbon number distribution of the product.

Tabak et al. suggested a reaction path for propylene conversion over ZSM-5 catalyst as shown in **Figure 1-2**. As it describes, four distinct reactions happen in the process, which are oligomerization of propylene to distinct oligomers, and then isomerisation and re-cracking to form a range of light olefins, and finally all the olefins re-polymerize to an equilibrium distribution of heavier olefins. Also their work also pointed out that for low carbon number products the isomer

distribution is approximately at equilibrium and the carbon number distribution is approximately at equilibrium.

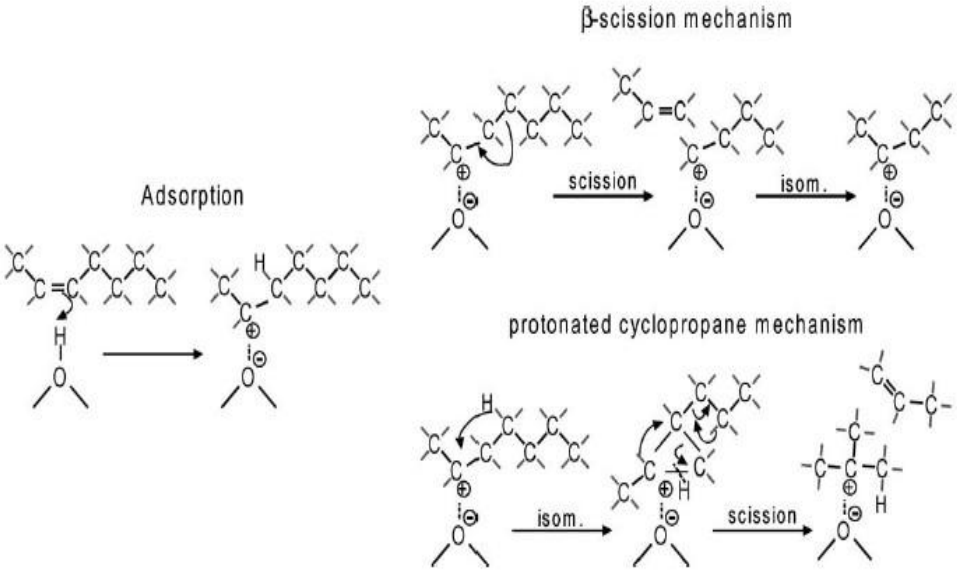


Figure 1-3. Cracking mechanisms illustrated by the reaction of 2-heptene (Hollander et al., 2002)

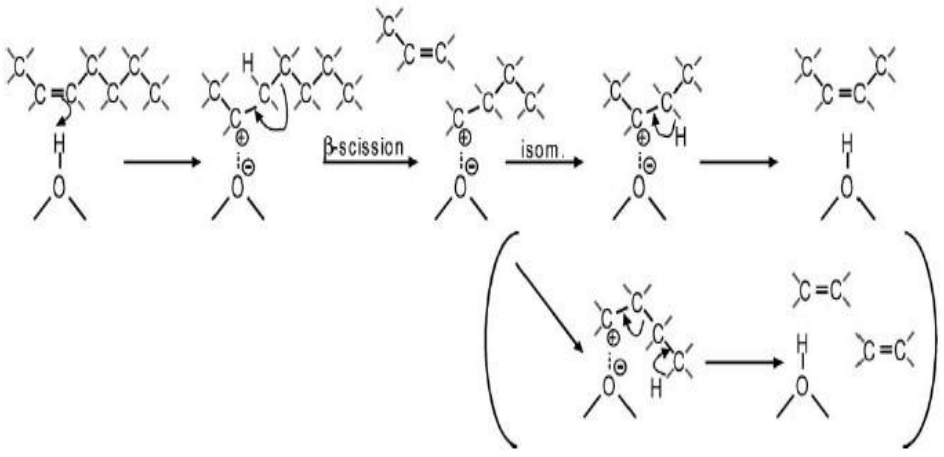


Figure 1-4. Monomolecular cracking mechanism over ZSM-5 illustrated by reaction of 2-heptene (Hollander et al., 2002)

One of the most important side reactions in the process is cracking of the heavier molecules. Cracking of heavier molecules will increase with an increase in reaction temperature and the decrease in pressure (Garwood, 1982, 1983; Tabak et al., 1983). The chemistry of cracking is shown in **Figure 1-3**, which is suggested by M.A. den Hollander et al. In general, the rate of cracking decreases in the order:

i-olefin> n-olefin> i-paraffin> n-paraffin > aromatics (Scherzer, 1989)

The cracking mechanism can be seen as chain reactions that involve the formation of intermediate of carbocations -- positively charged hydrocarbon species. The stability of carbocations decreases in the order of tertiary> secondary> primary (Scherzer, 1989). Reactions of olefins over acid catalysts mainly involve isomerization to more stable isomers and oligomerization to form larger molecules, which could further crack into smaller molecules (Hollander et al., 2002). At higher temperatures, monomolecular cracking, as shown in **Figure 1-4**, dominates the cracking process over the dimeric cracking mechanism, in which the carbenium ion absorbed on the catalyst acid site dimerizes with other hydrocarbons and can split-off a fragment to form a new molecule. They also conclude that the cracking of C_7 and larger molecules is restricted to simple β-scission reactions.

1.2.2 Spectroscopic analysis of hydrocarbons

The qualitative and quantitative analysis of the chemical components in the mixtures before and after the conversion of pyrolysis gas is one of most important and hardest sections in the process. Currently, the most commonly used analysis techniques to do this work are Gas Chromatography (GC) with single or multiple columns, usually combined with mass spectroscopy (MS) as GC-MS (Garcia-Perez et al., 2007; Singh et al., 2011; Xiu et al., 2006; Babich et al., 2011), Fourier Transform Infrared Spectroscopy(FTIR)/ Fourier Transform Near-Infrared Spectroscopy(FTNIR) (Boey et al., 2011; Ertas et al., 2010; Guo et al, 2010), Ultraviolet–Visible spectroscopy(UV-Vis) (Lu et al., 2011; Yin et

al., 2010; Patel et al., 2011; Li et al., 2008) and thermo-gravimetry (TG) (Girshrick et al., 1936, 1939; Hotelling, 1936, 1936).

Quick and accurate response, operational stability and easy-to-interpret are the main metrics a good online analysis system should have. But the current methods to interpret the spectroscopy are usually not favorable for online analysis and control, especially the GC-MS system and the TG method. GC can work online in most systems for the light chemicals, but still needs much preparation work such as calibration, and usually is expensive. Also GC is not efficient in analyzing heavy chemical components, and more, it is hard to be used for fast consecutive online tests. FTIR or FT-NIR is beneficial in fast online spectral analysis, but still needs manual processing when dealing with the overlapping in the spectral bands of more than one component in the chemical mixtures, and thus it is hard to be employed for automatic online control by itself. Especially for the conversion of pyrolysis gas system, the complexity of the components of the reaction product introduces more difficulties to perform online qualitative and quantitative analysis. And due to the dimensionality of the spectroscopic data, the analysis of difference among them also tends to be inefficient.

As stated above, the use of spectroscopy itself has some limitations or weaknesses in the area of online analysis and process control of pyrolysis gas conversion. So the target of our project is to use the spectroscopy combined with a description of the reaction network, as well as a mathematical and statistical method for pre-processing the data. Principal component analysis (PCA) will be employed to analyze the spectroscopy data from the reaction system. The spectroscopies that will be used in our project are mainly NIR-FTIR and UV-visible. Supplementary characterization of selected samples will be performed with GC-FID or GC-MS.

1.2.3 Lumped kinetic modelling

Lumped kinetic modelling as method has been long employed in kinetic modelling and study of complicated reaction networks (Reivich et al., 1985; Ayasse et al., 1997; Ancheya et al., 1999; Ranzi et al., 2001; Bollas et al., 2007; Grana et al., 2012). The similarity of reactants and reactions in the reaction network forms the basis of lumped modelling. In certain situations, lumping in a kinetic study is both necessary and practical. On the one hand, in certain situations the target chemical of research is certain type of chemicals rather than some single chemical, which makes it quite necessary to do a lumped kinetic study. On the other hand, due to the complexity of the reactions in the concerned processes and the large number of chemical species involved, it is not practical to focus on single chemical specie during kinetic modelling. Attention is more paid to certain types of chemicals or certain types of reactions. When spectroscopy is combined with the reaction network, the lumps must ideally be selected in such a way that a clear spectroscopic differentiation is also possible. For example, lumping olefins and paraffins of a specific carbon number may make sense for a chromatographic description where volatility can be a differentiator, but not for a spectroscopic description, where the functional group is the differentiator. With employing appropriate lumping algorithms according to the similarity of certain chemicals or certain types of reactions, similar or related chemical species are lumped together being treated as one component in the reaction network to simplify the whole reaction system. And within error allowed, appropriate lumped kinetic modelling can not only describe the abstract tendency of the special types of reactions, but it can also make accurate predictions of other parameters. For example, it can be used to predict the yield of chemical lumps, or concisely describe dependency of the product yield, or information on the reaction conditions, such as temperature and feed flow rate. For the development of a lumped kinetic modelling method, the number of lumps present in the published literature increased over the years as more

analysis information of reaction products can be gained. The intent with the increased complexity resulting from more detailed lumping is to build more complicated models for more accurate prediction. However, increased complexity does not necessarily lead to improved models and whatever the mathematical strategy, it must reflect the conversion chemistry fundamentals.

As stated above, lumping algorithm is employed in many reaction networks in oil refining research due to the complexity of the reaction system in oil chemistry. Examples of lumping strategies can be found for processes, such as fluid catalytic cracking (Shayegh et al., 2012; Fan et al., 2009; Bollas et al., 2007), catalytic pyrolysis of heavy-oil (Meng et al., 2011), hydrodesulphurization and hydrocracking of heavy residue oil (Elizalde et al., 2012), et al. Generally, there are two types of lumping algorithm classified by different grouping methods or targets. The first type is lumping on the basis of physical properties such as boiling point range, or other concerned parameters such as sulfur content, API gravity et al. This type of lumping often relies on empirical correlations and captures limited fundamental insight into the reaction chemistry. The second type of lumping is more focused on grouping of the chemicals according to the similarity of chemical structures or chemical reactions in the reaction network. This type of lumping is better suited for capturing knowledge about the reaction chemistry.

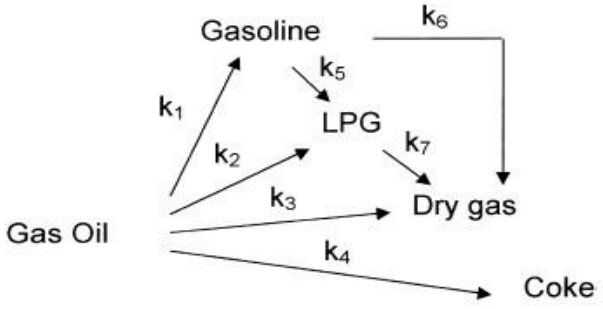


Figure 1-5. 5-lump kinetic model for catalytic cracking of gas-oil (Anchyta et al., 1999)

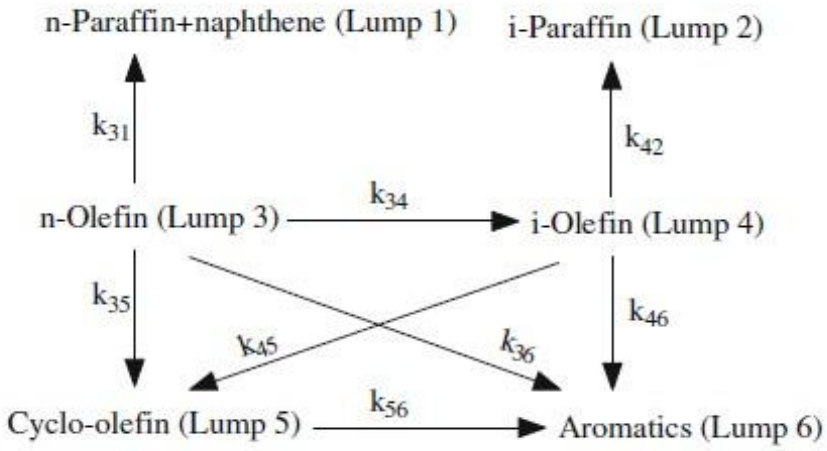


Figure 1-6. 6-lump kinetic model for FCC gasoline hydro-upgrading (Yu et al., 2009)

Lee et al., 1988 and 1989 grouped the reactants and products into four major lumps: gas oil, gasoline, light gas and coke, which is good example of the first type lumping. So did Ancheyta et al. 1999, by proposed a 5-lump kinetic model for catalytic cracking of gas-oil as shown in **Figure 1-5**. The 6-lump kinetic model built for FCC gasoline hydro-upgrading, as shown in **Figure 1-6**, is a typical lumped model of the second type as the lumps are grouped according to the similarity in chemical structure. The 7-lump model by Xu et al., 2006, and the 13-lump model for residual oil catalytic cracking by Meng et al. 2011 are also good examples of the second type of lumping.

In summary, both types of lumped modelling method have their advantages and disadvantages, the lumping based on physical properties is more valuable in predicting the concerned parameters of certain groups of chemicals separated accordingly, but could lead to loss of prediction accuracy due to its lack of theoretical support from the basic chemical reaction mechanism. And it is opposite for the lumping on the basis of chemical structures or similarity in mechanism of chemical reactions. Choosing which lumping method is for better

for the analysis of reaction system depends on the target of the modelling, the information that can be gained from the products, and practical realization of the mathematics.

1.3 Objective and Scope

The objective of this project was to develop online spectroscopy in combination with multivariate data analysis for the monitoring and control of dynamic systems. Of particular interest was the application of monitoring and control to pyrolysis gas conversion. For propose of studying conversion of pyrolysis gas to liquid product, our project used propylene as model feed. A continuous flow of propylene was converted over H-ZSM-5 catalyst in a packed bed reactor. The effect of reaction conditions on the yields of chemicals, and the kinetics of the reaction network were studied and modelled. Multivariate data analysis was applied to analyze the offline and online chromatographic and spectroscopic data for understanding the process, and to optimize the modelling and process control of the reaction.

This study first introduces the project background and an associated literature review in this chapter. Multivariate data analysis including Principle Components Analysis (PCA) and Hierarchical Clustering Analysis (HCA), is introduced in Chapter 2. Chapter 3 mainly presents the experimental procedures and methods of this study. Chapter 4 discusses the effects of space velocity and reaction temperature on the yields of chemical product in the catalytic conversion process and in Chapter 5, a lumped kinetic model is built for analysis of the reaction network and for practical prediction of the yield of chemical lumps under various reaction conditions of space velocity and reaction temperature. PCA and HCA are applied in Chapter 6 to analyze the reaction data for analysis and control of the process and certain strategies are proposed for future process analysis and control. Chapter 7 sums up the conclusions of this study and future work is also proposed for further study.

References:

J. Anchyta, F. Lopez-Isunza, E.Aguilar-Rodriguez.5-lumped model for gas oil catalytic cracking. Applied Catalysis A: General, 1999, 177(2):227-235.

A.R. Ayasse, H. Nagaishi, E.W. Chan. Lumped kinetics of hydrocracking of bitumen. Fuel, 1997, 76(11):1025-1033.

I.V. Babich, M. van der Hulst, L. Lefferts, et al. Catalytic pyrolysis of microalgae to high-quality liquid bio-fuels. Biomass and Bioenergy, 2011, 35(7):3199-3207.

C. Baerlocher, L. B. McCusker, D. H. Olson. Atlas of zeolite structure types; http://www.iza-sc.ethz.ch/IZA-SC/Atlas/AtlasHome.html, 2000.

R. Bentley, R. Miller, S. Wheeler, et al. UKERC Review of Evidence for Global Oil Depletion--Technical Report 7: Comparison of Global Oil Supply Forecasts. UK Energy Research Council, 2009.

P. Boey, M.I. Saleh, N. Sapawe, et al. Pyrolysis of residual palm oil in spent bleaching clay by modified tubular furnace and analysis of the products by GC-MS. Journal of Analytical and Applied Pyrolysis, 2011, 91(1):199-204.

G.M. Bollas, A.A. Lappas, D.K.Iatridis, et al. Five-lump kinetic model with selective catalyst deactivation for the prediction of the product selectivity in the fluid catalytic cracking process. Catalysis Today, 2007, 127(1-4):31-43.

C. De Castro, L.J. Miguel, M. Mediavilla. The role of non-conventional oil in the attenuation of peak oil. Energy Policy, 2009, 37(5):1825-1833.

I. Elizalde, J. Ancheyta. Modeling the simultaneous hydrodesulfurization and hydrocracking of heavy Residue Oil by using the continuous kinetic lumping approach. Energy & Fuels, 2012, 26(4):1999-2004.

M. Ertas, M.H. Alma. Pyrolysis of laurel (Laurus nobilis L.) extraction residues in a fixed-bed reactor: Characterization of bio-oil and bio-char. Journal of Analytical and Applied Pyrolysis, 2010, 88(1):22-29.

Y. Fan, J. Yin, G. Shi, et al. A six-lump kinetic model for olefin hydrogenation, hydroisomerisation and aromatization in FCC gasoline hydro-upgrading. Catalysis Letters, 2009, 129(1-2):181-188.

M. Garcia-Perez, A. Chaala, H. Pakdel, et al. Vacuum pyrolysis of softwood and hardwood biomass- Comparison between product yields and bio-mass properties. Journal of Analytical and Applied Pyrolysis, 2007, 78(1):104-116.

W.E. Garwood. Conversion of C2 to C10 olefins into higher olefins on the synthetic zeolites ZSM-5. Erdol & Kohle Erdgas Petrochemie, 1982, 10(35): 479-497.

M.A. Girshick. Principal components. Journal of American Statistical Association, 1936, 31(193):519-528.

R. Grana, A. Frassodalti, C. Saggese, et al. A wide range kinetic modeling study of pyrolysis and oxidation of methyl butanoate and methyl decanoate – Note II: Lumped kinetic model of decomposition and combustion of methyl esters up to methyl decanoate. Combustion and Flame, 2012, 159(7):2280-2294.

X. Guo, S. Wang, Z. Guo, et al. Pyrolysis characteristics of bio-oil fractions separated by molecular distillation. Applied Energy, 2010, 87(9):2892-2898.

M.A. den Hollander, M. Wissink, M. Makkee, et al. Gasoline conversion: reactivity towards cracking with equilibrated FCC and ZSM-5 catalysts. Applied Catalysis A: General, 2002, 223(1-2):85-102.

H. Hotelling. Relations between two sets of variates. Biometrika, 1936, 28:321-377.

H. Hotelling. Simplified calculation of principal components. Psychometrika. 1936, 1(1):28-36.

L. Hughe and J. Rudolph. Future world oil production: growth, plateau, or peak? Current Opinion in Environmental Sustainability, 2011, 3(4):225-234.

IEA: Key World Energy Statistics 2010. Paris: International Energy Agency, 2010.

V. N. Ipatieff, B.B. Corson, G. Egloff. Polymerization, a new source of gasoline. Industrial and Engineering Chemistry, 1935, 27:1077-1081.

G.T. Kokotalio, S.L. Lawton, D.H. Olson, et al. Structure of synthetic zeolites ZSM-5. Nature, 1978, 5652(272):437-438.

J. Li, L. Wu, Z. Yang. Analysis and upgrading of bio-petroleum from biomass by direct deoxy-liquefaction. Journal of Analytical and Applied Pyrolysis, 2008, 81(2):199-204.

R. Lu, G.P. Sheng, Y.Y. Hu, et al. Fractional characterization of a bio-oil derived from rice husk. Biomass and Bioenergy. 2011, 35, 671-678.

X. Meng, C Xu, L. Li, et al. Kinetics of catalytic pyrolysis of heavy gas oil derived from Canadian synthetic crude oil. Energy& Fuels, 2011, 25(8):3400-3407.

P. Murugan, T. Mani, N. Mahinpey and M. Dong. Pyrolysis kinetics of athabasca bitumen using a TGA under the influence of reservoir sand. Canadian Journal of Chemical Engineering, 2012, 90(2):315-319.

A. G. Oblad, G. A. Mills, H. Heinemann. Catalysis IV: Emmert P.H.Ed. Reinhold, New York, 1958, Ch. 4: 341-345.

R.N. Patel, S. Bandyopadhyay, A. Ganesh. Extraction of cardanol and phenol from bio-oils obtained through vacuum pyrolysis of biomass using supercritical fluid extraction. Energy, 2011, 36(3):1535-1542.

E. Ranzi et al., Lumping procedures in detailed kinetic modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures. Progress in Energy and Combustion Science. 2001, 27:99-139.

M. Reivich, A. Alavi, A. Wolf, et al. Glucose metabolic-rate kinetic-model parameter determination in Humans- the lumped constants and rate constatns for [F-18] fluorodeoxyglucose and [C-11] deoxyglucose. Journal of Cerebral Blood Flow and Metabolism, 1985, 5(2):179-192.

J. Scherzer. Octane-enhancing, zeolitic FCC catalysts: scientific and technical aspects. Catalyst Review-Science and Engineering, 1989, 31(3):215-354.

F. Shayegh, A. Farshi, A. Dehgan. A kinetics lumped model for VGO catalytic cracking in a fluidized bed reactor. Petroleum Science and Technology, 2012, 30(9):945-957.

R.K. Singh, K.P. Shadangi. Liquid fuel from castor seeds by pyrolysis. Fuel, 2011, 90(7):2538-2544.

S. A. Tabak, F. J. Krabeck, W. E. Garwood. Conversion of propylene and butylene over ZSM-5 catalyst. AIChE Journal, 1983, 32:1526-1531.

S. Xiu, A. Shahbazi, V. Shirley, et al. Hydrothermal pyrolysis of swine manure to bio-oil: Effects of operating parameters on products yield and characterization of bio-oil. Journal of Analytical and Applied Pyrolysis. 2010, 88(1):73-79.

O. Xu, H. Su, S. Mu, et al. 7-lump kinetic model for residual oil catalytic cracking. Journal of Zhejiang Univ Science A, 2006, 7(11):1932-1941.

S. Yin, R. Dolan, M. Harris, et al. Subcritical hydrothermal liquefaction of cattle manure to bio-oil: Effects of conversion parameters on bio-oil yield and characterization of bio-oil. Bioresource Technology, 2010, 101(10):3657-3664.

Chapter 2

Multivariate Data Analysis: Principal Component Analysis and Hierarchical Clustering Analysis

2.1 Introduction

Based on the statistical principal of multivariate statistics, multivariate data analysis involves dealing with more than one variable at a time, and usually the dimension of data analyzed is much higher than one. With statistical analysis, this technique takes variables of interest into account to make inferences of the involved process or phenomena of the data source.

Principal Component Analysis (PCA) and Hierarchical Clustering Analysis (HCA) are two of the most widely used multivariate data analysis methods. PCA and HCA are applied in many areas of science and engineering on data dimension reduction and similarity probing to discover the underlying important and comprehensive meaning of the data. They are also the two data analysis methods employed in the data analysis of the work of this study. PCA and HCA are employed to study the reaction process, both to analyze the effect of reaction conditions on the reaction system and the properties of the product, and to analyze of the similarity of the product samples and the reaction conditions, to help to understand the reaction process. This chapter gives a brief introduction to PCA and HCA.

2.2 Principal Component Analysis

2.2.1 Introduction

Introduced by Karl Pearson in 1901, Principal Component Analysis (PCA) has been one of the most used multivariate high-dimensional data analysis methodologies in data exploration and modelling. With continuous development for over a century, PCA is widely applied in diverse areas such as engineering data analysis and statistics (Lavinea et al., 2012; Dobos et al., 2012; Lin, 2012), pattern recognition (Eiceman et al., 2006; Lavine et al., 2001; Ozawa et al., 2005; Karasinski et al., 2007), life science (Karasinski et al., 2007; Lee et al., 2012; Yang et al., 2011), and meteorology (Smith et al., 2012; Gorgas et al., 2012; Cheng et al., 2012). PCA is also commonly employed as data pre-processing for subsequent analysis such as clustering analysis, which is introduced in a later section.

"The central idea of principal components analysis is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables." (Jolliffe, 2002)

Being considered one of the most used and practical mathematical procedures in high-dimensional data analysis and pattern recognition, PCA mainly focuses on converting a set of observations of data into a set of linearly uncorrelated variables by a mathematical method named orthogonal transformation, which can efficiently cut the redundancy of the dataset with further data regression treatment. The generated new variables are known as principal components with the first principal component accounting for the largest possible variance, followed the second principal component accounting for the second most possible variance, and the same mathematic meaning applies for the succeeding principal components. Obviously, principal components are mathematically independent only if the data set is jointly normally distributed. Loading from each principal component on the total data variance which

represents the contribution of this principal component to the whole variance of the data matrix is calculated in PCA. Generally, after discarding the later principal components with loading or accumulated loading beyond a certain threshold value, the dimension of the data set could be greatly reduced without losing much information with the retained principal components. To a degree, PCA can artfully reveal the internal structure of the dataset in a way which best explains the variance of the data set and simplify the intuitive view of the whole data set. Graphically, PCA extracts dimensional directions of the high-dimension data where the data are more extended, and it proceeds one by one along the directions until the last dimensional direction is represented with the least variance of the whole data set. Each dimensional direction represents one principal component and its corresponding eigenvector, and theoretically the dimensional directions are independent and orthogonal to each other.

Figure 2-1 gives one of the simplest examples of PCA in which two maximum variation directions are calculated from the original data, with variation in the first principal component (PC1) direction dominating the variation in the second principal component (PC2) direction.

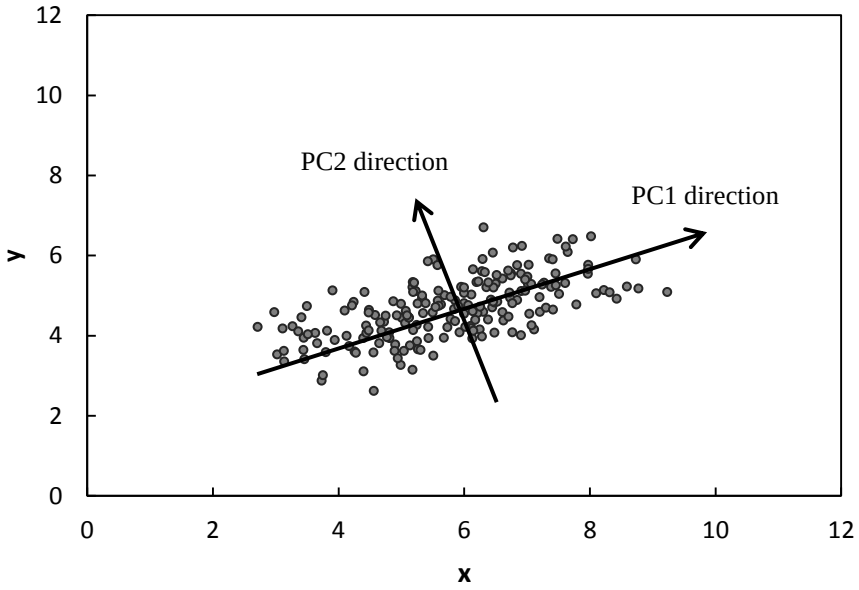


Figure 2-1. Principal component analysis of a data set in two variables (*x*, *y*). The data set obeys a normal distribution centered at (6, 4.8), with a standard deviation roughly of 1.553 in the direction of (1, 0.4) as PC1 direction shows, and deviation of 0.621 in the orthogonal direction of PC2

In reality, the dimension of the achieved principal components in a PCA is exactly the same with the dimension of variables in the original data set. But only a few principal components may be used for further analysis or be employed as criterion variables since these principal components are enough to interpret most of the variance present in the original data set. Thus with certain regression processing on the principal components, the dimension of the original data set can be reduced efficiently.

2.2.2 Methodology of Principal Components Analysis

As stated earlier, the result of PCA is a set of principal components. In the set of principal components the first principal component accounts for a maximal fraction of the total variance in the observed variables. Under typical conditions, this means the first component will be correlated with at least, normally much more than, two of the observed variables as long as there are correlations

between the variables. The same applies to subsequent principal components: the second principal component accounts for the maximal variance that is less than and not included in the first principal component. The mathematical detail of the implementation of PCA is provided below.

Define a $m \times n$ data matrix O^T, with o_1, o_2, …, o_n being its corresponding $m \times 1$ row vectors. In this data matrix, each of the rows represents a different experiment, and each of the columns represents a distinct output (measured) variable (e.g., concentration, temperature, etc).

$$O^T = [o_1, o_2, ..., o_n]$$

Usually, the first step is to standardize and center the data, which involves subtracting the empirical mean from the original data to obtain a new data matrix X^T with the same dimension, but with zero empirical mean as shown below:

$$\bar{o} = \frac{1}{n} \cdot \sum_{i=1}^{i=n} o_i$$

$$X^T = O^T - \bar{o}$$

The next step is computing the covariance matrix (Cov) of X^T after centering the data,

$$Cov = X^T X = \frac{1}{m} \cdot \sum_{i=1}^{i=m} x_i^T x_i$$

The singular value decomposition of X^T is given by

$$X^T = W \Sigma V^T,$$

where the $m \times m$ matrix W is the matrix of eigenvectors of the covariance matrix $X^T X$, and the $n \times n$ matrix V is the matrix of eigenvectors of the covariance matrix XX^T. The PCA that preserves dimensionality is then given by:

$$Y^T = XW = V \Sigma^T W^T W = V \Sigma^T .$$

As calculated above, a reduced dimensional representation can be gained by projection of X^T down to the reduced space defined by only the first L eigenvectors W_L, then

$$Y = W_L^T X^T = \Sigma_L V^T,$$

where $\Sigma_L = I_{L\times m}$ with $I_{L\times m}$ the $L \times m$ rectangular identity matrix.

As we get the matrix V, the next step is computing the corresponding eigenvalues λ_1, λ_2, …, λ_n,

$$\begin{bmatrix} \lambda_1 & ... & 0 \\ ... & ... & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} = V^{-1} Cov V$$

Then, we calculate the eigenvectors e_1, e_2,…, e_n, which form a set that in turn form the basis for the new projected coordinate based on this new coordinate, the corresponding eigenvalues account for the variance in the direction of each eigenvector.

Obviously, it can be seen that PCA is sensitive to the relative scaling of the original data as it concerns the scaling of the covariance or correlation matrix. The first principal component represents the largest variance of the data set and the succeeding components represent diminishing variance.

Numerically, principal components are more like the scores of different variance directions of the source data; the larger the score (principal component), the larger the variance it presents in that direction. From a mathematical point of view, one principal component could be responsible for several correlated variables in the original data source, thus enabling dimension reduction of the original data without losing much information.

As the principal components are calculated, data projection can be done to extract the main meaningful variables in the original data. Since the first couple of principal components usually account for most of the variance, it is thus possible to explore the most important information of the variance of the whole set data.

On the basis of eigenvalue analysis, PCA can be easily done by eigenvalue-decomposition of a data covariance or correlation matrix, and also can be achieved by singular value decomposition of a data matrix, usually with data pre-processing such as data mean centering and standardization.

To reiterate the essence of PCA, with orthogonal linear transformation, PCA transforms the data into a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate which is the first principal component, and the second coordinate accounts for the second principal component, or the second greatest variance, and so on. Below are the mathematical details of PCA.

This brings us to the critical step that practically implements dimensional reduction, which in fact is choosing the j largest eigenvalues and then projecting back to the original dataset to get a lower-dimension dataset.

2.2.3 Information loss

In PCA, dimension reduction (i.e. retaining only a limited number of principal components) will introduce certain loss of information of the original data set. This information loss can be minimized as needed. To make sure the information loss is in an acceptable range, usually a certain threshold value is chosen to determine how many principal components will be retained and the last retained principal component with index j could be determined as below,

$$\frac{\sum_{i=1}^{i=k} \lambda_i}{\sum_{i=1}^{i=n} \lambda_i} \geq threshold$$

The threshold value here could be 90% or 95% or any value between 0 and 1 as needed. The larger the value of the threshold is, the less is the loss of information from the original dataset. Generally this also means the more principal components will be retained and less data dimensions reduced.

2.2.4 Application in this study

This study focuses on the using of offline chromatography and online spectroscopy for data analysis. The aim is to provide a link between the analytical input such as reaction conditions and the reaction network, kinetic modeling, process analysis and control of the process for pyrolysis gas conversion to liquid. Since there are so many factors that affect the reaction system, such as the feed flow rate, product flow rate, conversion of feed gas,

reaction temperature and so on, a reduction in data dimensionality is needed. PCA is employed to analyze the whole set of parameters and variables including the spectra generated in the experiments. This could greatly help to identify the main information from the complex system in which there probably is a quite strong potential for correlation between different parameters.

2.3 Hierarchical Clustering Analysis

2.3.1 Introduction

Clustering analysis or clustering is a common technique of multivariate data analysis in many fields of science and engineering, including image analysis (Tang et al., 2012; Hugo et al., 2008), machine learning (Fiser et al., 2012; Fan et al., 2008), and pattern recognition (Mok et al., 2012; Hugo et al., 2008). By assigning or grouping a set of objects into groups or clusters (under evaluation of certain criterions), clustering aims to explore and exploit the relative greater similarity of concerned objects to each other in the same cluster than to those in other clusters.

The meaning of similarity of objects in the same cluster in clustering analysis depends on the algorithm employed in a particular situation. Even implemented on the same data set, the clusters grouped by different algorithms could vary significantly in the properties, cluster structure and meaning of the concerned similarity. These algorithms includes connectivity based clustering, centroid-based clustering, distribution-based clustering and density-based clustering, to name a few.

The algorithm employed in this study is connectivity based clustering, also called hierarchical clustering. Hierarchical clustering is based on the difference metric (the distance of different observations or sub-clusters in a high dimension space, one simple example is the Euclidean distance between two points in a 2-dimension space) which represents the dissimilarity between the concerned observations or sub-clusters. Hierarchical clustering analysis provides an extensive hierarchy of clusters that merge with each other at certain distance (measure of difference) by the chosen algorithm of distance calculation.

2.3.2 Basics and Mathematics

HCA is further classified into two types, namely agglomerative HCA and divisive HCA. Agglomerative HCA takes each object as a separate cluster in the first step, and step by step, two clusters are merged together to form a higher-level hierarchy. This is actually an exact "bottom-up" strategy. Divisive HCA employs "top-down" method, as it takes all the observations as one cluster in the first step and divides each cluster into two, step by step, as it is moving down the hierarchy. Due to the difference in mathematic complexity between the two HCA approaches, agglomerative HCA that is easier to perform is more often employed in data analysis.

As distance has been chosen as criteria for merging clusters, an appropriate measure of distance of clusters and the linkage criterion are needed to be specified to determine the dissimilarity of clusters when performing HCA.

As shown in **Table 2-1**, the most often used metrics for HCA are Euclidean distance, square Euclidean distance, Manhattan distance and maximum distance.

Table 2-1. Metrics commonly used in HCA

Metric Names	*Formula for distance calculation*
Euclidean distance	$dis(x,y)=\sqrt{\sum_i (x_i-y_i)^2}$
Square Euclidean distance	$dis(x,y)=\sum_i (x_i-y_i)^2$
Manhattan distance	$dis(x,y)=\sum_i \lvert x_i-y_i \rvert$
Maximum distance	$dis(x,y)=\max_i \lvert x_i-y_i \rvert$

Linkage algorithm is the algorithm that determines the criteria that used for constructing the cluster structures. The distances between the existed clusters are calculated according to the metric method, and two clusters are chosen according

to the chosen linkage method merge to form a new cluster. And the most used linkage algorithms are listed in **Table 2-2**.

Table 2-2. Linkage methods commonly used in HCA

Linkage Names	*Formula for linkage calculation*
maximum linkage	$\max\{dis(x,y): x \in clusterX, y \in clusterY\}$
minimum linkage	$\min\{dis(x,y): x \in clusterX, y \in clusterY\}$
mean linkage	$\frac{1}{n_X n_Y}\sum_{x\in X}\sum_{y\in Y} dis(x,y)$
inner square linkage	$\min\left[\frac{1}{n_Z}\sum_{i,j}(z_i - z_j)^2 - \frac{1}{n_X}\sum_{i,j}(x_i - x_j)^2\right], Z = X \cup Y$

With the chosen metric and linkage criteria, HCA can be easily performed by MATLAB programming with the corresponding functions *pdist( ), linkage( ), dendrogram( )* and so on in its toolbox.

2.3.3 Application in this study

The application of HCA in is study deals mainly with three aspects:

First, HCA is employed to explore the relative similarity of samples by analysis of the sample spectra. Online or offline spectra data contains lots of information about the chemical components of the product samples. Differences are usually hard to be detected visually due to the overlapping of spectral absorption, or incomplete chromatographic separation of different compounds. These similarities can be detected by HCA of the raw data of the spectra of the concerned samples.

Second, HCA is applied to analyze the comprehensive effects from several reaction parameters on the reactions. Effects from a single parameter, such as temperature or feed flow rate, are easy to detect by a series of experiments, but for changes of several parameters at the same time, it is harder to determine how

they affect the reactions. With HCA, it is easy to find the comprehensive effect on the reactions from several parameters, and also from the similarity of product samples. In this way it is easy to determine whether it is a cooperative effect, or effect of the offset one another between different reaction parameters on the same reaction.

Third and the last, HCA is used to check the outliers in the process control of the reaction system. The idea for this is that, with the analysis results of the first and the second application of HCA described above, a mapping of values of reaction parameters to the properties of products under these conditions can be built. This mapping can be used in process control. Any serious deviation from this mapping will give a warning to the signal system that the data is in error or abnormal.

References:

K. Pearson. On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine, 1901, 2(6):559–572.

B.K. Lavinea, K. Nugurua, N. Mirjankara, J. Workmanb. Development of carboxylic acid search prefilters for spectral library matching. Microchemical Journal, 2012, 103:21-36.

L. Dobos, J. Abonyi. On-line detection of homogeneous operation ranges by dynamic principal component analysis based time-series segmentation. Chemical Engineering Science, 2012, 75(18):96-105.

J. Lin. A novel design of wafer yield model for semiconductor using a GMDH polynomial and principal component analysis. Expert Systems with Applications, 2012, 39(8):6665-6671.

G.A. Eiceman, M. Wang, S. Pradad, et al. Pattern recognition analysis of differential mobility spectra. Analytic Chimi Acta, 2006, 579:1–10.

B.K. Lavine, C.E. Davidson, A.J. Moores. Genetic algorithms for spectral pattern recognition. Vibrational Spectroscopy, 2008, 28(1):83–95.

D. Lee, H. Seung. Learning the parts of objects by non-negative matrix factorization. Nature, 401 (6755) (1999), 788–791.

S. Ozawa, S. Toh, S. Abe, et al. Incremental learning of feature space and classifier for face recognition. IEEE Transactions on Neural Networks, 2005, 18(5-6):575–584.

J. Karasinski, L. White, Y. Zhang, et al. Detection and identification of bacteria using antibiotic susceptibility and a multi-array electrochemical sensor with pattern recognition. Biosensors and Bioelectronics, 2007, 22(11):2643-2649.

S. Lee, P.E. Michael, R. Duncan, et al. Sparse Principal Component Analysis for Identifying Ancestry-Informative Markers in Genome-Wide Association Studies. Genetic Epidemiology, 2012, 36(4):293-302.

X. Yang, S. Gao, S. Xu, et al. Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. Molecular Breeding, 2011, 28(4):511-526.

G.L. Smith, P.E. Mlynczak, G.L. Potter. A technique using principal component analysis to compare seasonal cycles of Earth radiation from CERES and model computations. Journal of Geophysical Research-atomospheres, 2012, 117: D09116.

T. Gorgas, M. Dorninger. Concepts for a pattern-oriented analysis ensemble based on observational uncertainties. Quarterly Journal of the Royal Meteorological Society, 2012, 138(664):769-784.

I. Cheng, L. Zhang, P. Blanchard, et al. Source-receptor relationships for speciated atmospheric mercury at the remote Experimental Lakes Area, northwestern Onoilio, Canada. Atmospheric Chemistry and Physics, 2012, 12(4):1903-1922.

I.T. Jolliffe. Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002. ISBN 978-0-387-95442-4.

J. Tang, B. Jiang, C. Chang, et al. Graph structure analysis based on complex network. Digital Signal Processing, 2012, 22(5):715:725.

Z. Hugo, A. Chritophe, M. Vincent. Fast online graph clustering via Erdos-Renyi mixture. Pattern Recognition, 2008, 41(12):3592-3599.

K. Fiser, T. Sieger, A. Schumich, et al. MRD Monitoring of Childhood ALL Using Hierarchical Clustering and Support Vector Machine Learning of Complex Multi-Parameter Flow Cytometry Data. Blood, 2008, 112(11):536-536.

C.Y. Fan, P.S. Fan, T.Y. Chan, et al. Using hybrid data mining and machine learning clustering analysis to predict the turnover rate for technology professionals. Expert Systems with Applications, 2012, 39(10):8844-8851.

P.Y. Mok, H.Q. Huang, Y.L. kwok, et al. A robust adaptive clustering analysis method for automatic identification of clusters. Pattern Recognition, 2012, 45(8):3017-3033.

Chapter 3

Experimental Design and Procedure

3.1 Chemicals and equipment

The main equipment employed in this study for the experiment and the analyses are listed in Table 3-1 as below.

Table 3-1. Equipments employed

Equipment	Model	Supplier
Gas Chromatograph with Flame Ionization Detector and Thermal Conductivity Detector (GC-FID-TCD) for gas sample	7890A	Agilent Technologies Inc.
GC-FID-TCD for liquid sample	7890A	Agilent Technologies Inc.
Gas Chromatograph with Mass Spectroscopy for liquid sample	7890A	Agilent Technologies Inc.
Thermometer	54-2-B	Fluke Corporation.
Thermocouple	N.A.*	Omega Engineering Inc.
Refrigerated-Heating Circulators	F25-HE	Julabo Company.
Gas flowmeter	AMD 1000	Agilent Technologies Inc.
FTIR spectrometer	MB3000	ABB Group.
NIR spectrometer	RW-NIR-SR	StellerNet Inc.
UV-Vis spectrometer	Black-Comet	StellerNet Inc.
Liquid spectra flow cell	TFC-S13-3	Harrick Scientific Inc.
Tube Furnace	N.A.*	Thermo-Scientific Inc.
Steel tubes & fittings	N.A.*	Swagelok Company.
Mass flow controllers	N.A.*	Brooks Automation, Inc.

*N.A: Not Applicable.

The main chemicals used for experiment and analysis in this study are shown in **Table 3-2**. And the chemicals used for calibration of GC-FID-TCD are listed Table E-1 in Appendix X.

Table 3-2. Chemicals used in the experiment and analysis in this study

Chemicals	**Purity** (wt %)	**Supplier**
ZSM-5 catalyst	≥99.9%	Zeolyst International Inc.
Silicon Carbide	≥ 99.0%	Kramer Industries Inc.
Carbon sulfide for GC-MS	≥ 99.5%(GC)	Sigma-Aldrich Corporation.
Propylene	≥ 99.5%	Praxair Canada Inc.
Hydrogen for GC	≥ 99.99%	Praxair Canada Inc.
Nitrogen	≥ 99.99%	Praxair Canada Inc.
Helium for GC	≥ 99.99%	Praxair Canada Inc.
Compressed Air	≥ 99.99%	Praxair Canada Inc.

3.2 Catalyst and reactor design

ZSM-5 zeolite catalyst from Zeolyst International Inc. was employed for the catalysis of the reaction. The properties of this catalyst provided by the supplier are shown in Table 3-3. The catalyst was blended with 400-mesh Silicon carbide (SiC) from Karmer Industries Inc. for use in the reactor. The catalyst was activated by calcining the zeolite powder at 550 ℃ for 5 hours. The blend ratios of zeolites catalyst to silicon carbide of 1:5, 1:10 and 1:20 were tested in the first round of experiment. With considering the strong exothermicity of the process which makes it hard to measure and control the reaction temperature under high blending ratio of catalyst to silicon carbide of 1:5 and 1:10, the blending ratio of catalyst to silicon carbide of 1:20 is employed for the process in this study under which condition all the data in this study is collected for analysis.

Table 3-3. Property of ZSM-5 catalyst by the supplier

Property	Value
SiO_2/Al_2O_3 Mole Ratio	50
Nominal Cation Form	Ammonium
Na_2O Weight %	0.05
Surface Area, m^2/g	425
Specific gravity	>1
Solubility in water	Negligible

The configuration of packed reactor is shown in **Figure 3-1**. The zeolite catalyst is diluted by 400-mesh SiC powder to reduce the thermal effect of the reactions. Other sizes of SiC are also used to pack the reactor to get smooth and continuous tubular axial reaction flow and to constrain the radial flow, which is critical for the kinetic model built in Chapter 4 of the study.

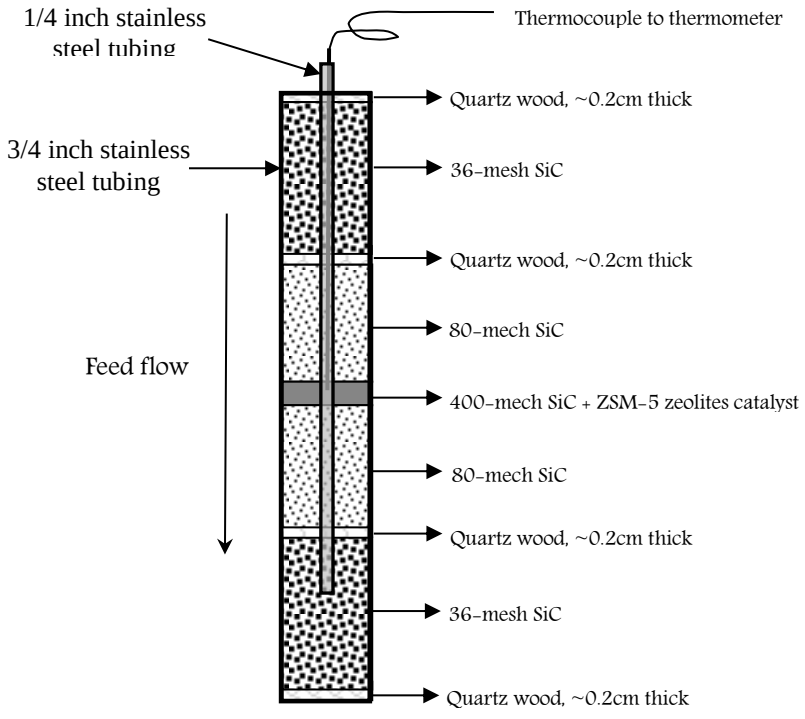


Figure 3-1. Packing of the reactor for continuous reaction flow

As shown in **Figure 3-1**, a stainless-steel tube with size of 3/4 inch and length of 34cm is used as reactor for the process. To make the thermocouple free to move inside the reactor for measurement of axial temperature profile, a steel tube with size of 1/4 inch and length of 30cm is fixed along the axis of the reactor with one side sealed and the other side open for the thermocouple. A T-junction is used for the inlet of the reactor to connect both the feed flow in tube and the tube for the thermocouple as shown in **Figure 3-2**.

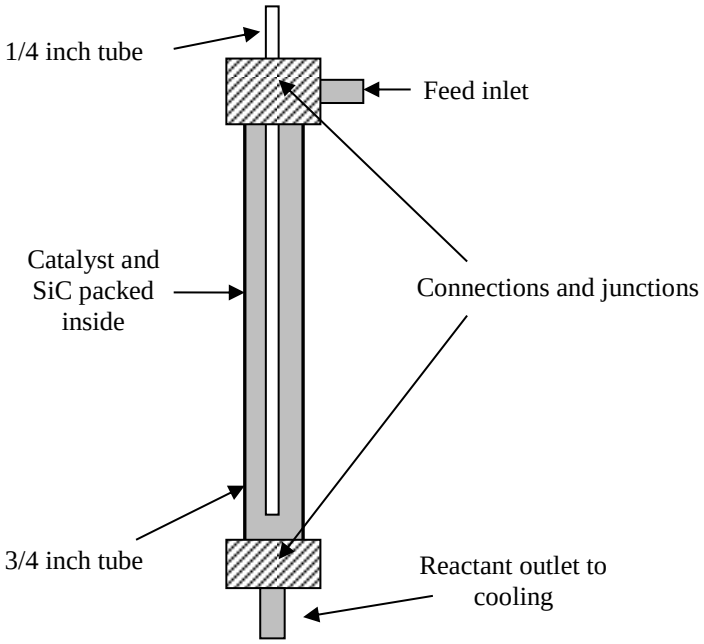


Figure 3-2. Sketch configuration of the reactor

3.3 Experimental design and setup

A continuous tubing reactor system was built for the reaction of propylene conversion, and the reactor was packed with zeolite catalyst, silicon carbide and quartz wool as shown in **Figure 3-1**.

As shown in **Figure 3-3**, the feed gas flowed into the tubing reactor with its flow rate controlled by the gas flow controller. The generated products flowed out and were cooled down indirectly by the continuous chilled water flow from the refrigerated-heating circulator as shown in **Figure 3-4**. After being cooled the products were separated into gas and liquid and were collected by liquid

container and gas bag individually for characterization. The temperature of the gas flow after separation at the outlet was also measured for calculation of the mass flow the gas products.

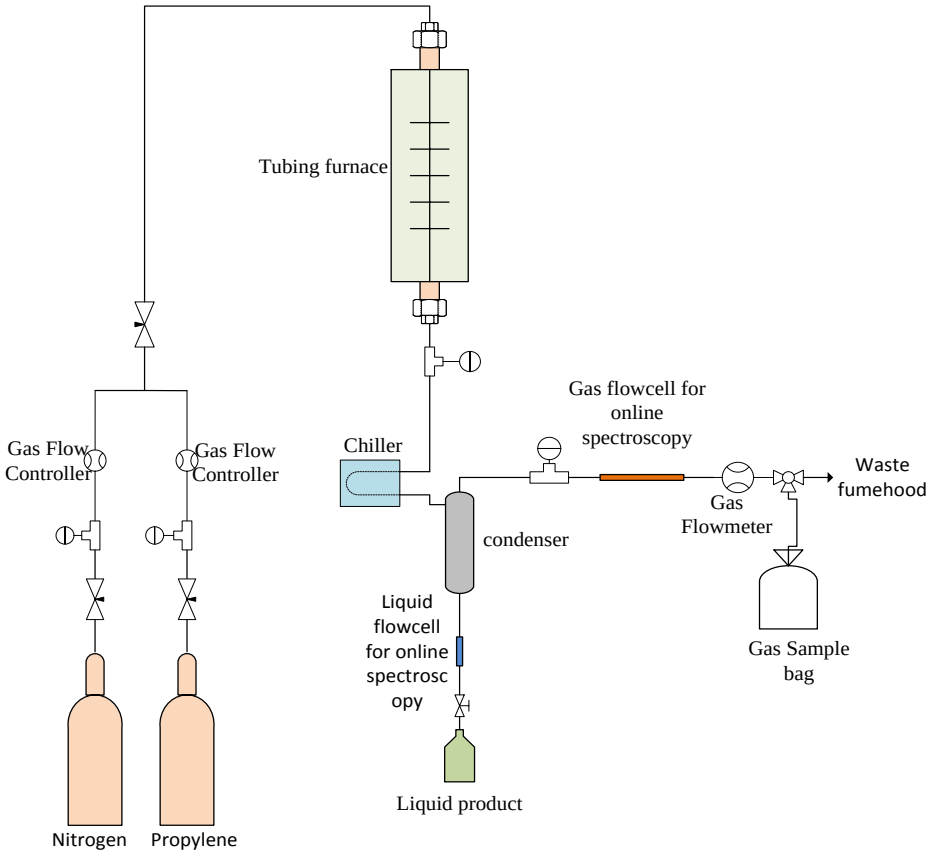


Figure 3-3. Sketch implementation of the reaction system

Online flow cells are also installed on the path of liquid product flow and gas product flow for online spectroscopy as shown in **Figure 3-4**. For each flow cell installed, one end is connected to a light source and the other end is connected to a light sensor for transmission spectra detection. And the 13×2mm Potassium bromide (KBr) window is used for all the light paths.

For details as shown in **Figure 3-4**, the liquid products after condensation first flows between two windows through a flow cell with 5cm optic path length for UV-Vis spectroscopy detection, following which a flow cell with 0.2cm light path for FTIR spectra measurement is also installed as shown in **Figure 3-4**. The

gas products after being cooled flows through a 50cm long tubing cell IR and NIR-Vis spectra detection.

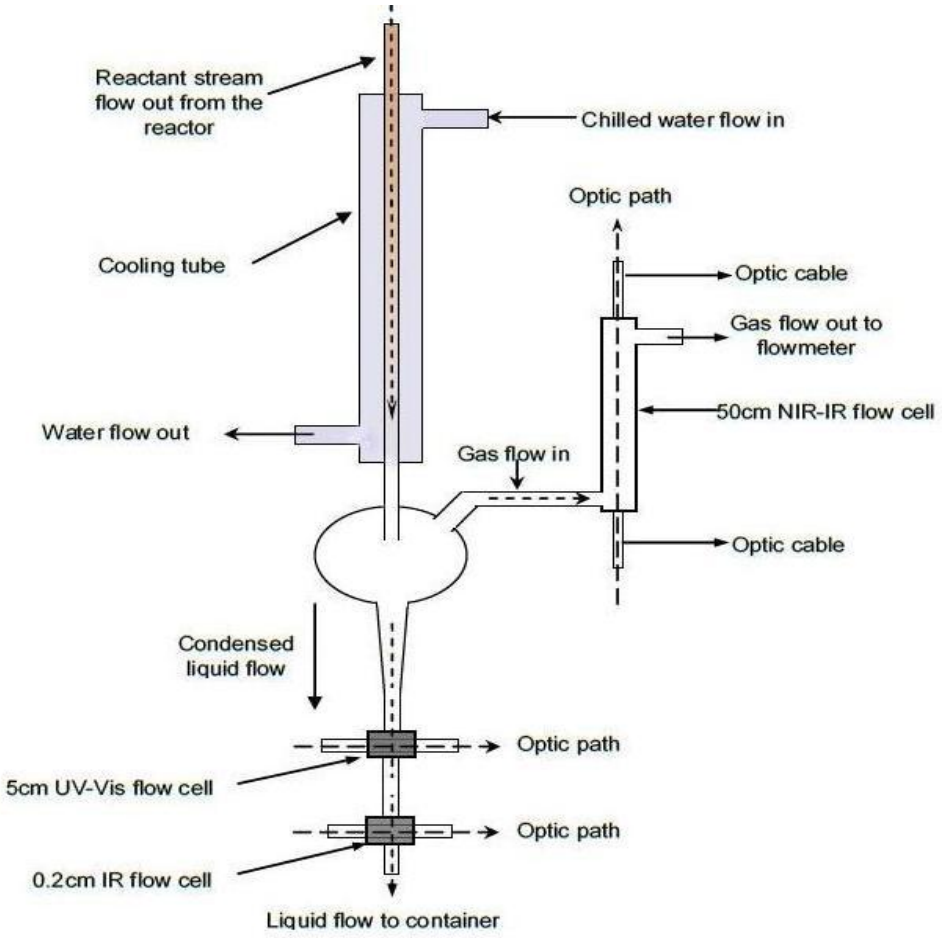


Figure 3-4. Sketch of the products condensation and the online flow cells

3.4 Method and analysis

3.4.1 Method

Figure 3-5 shows the sketch of the method and the experimental procedure of this study.

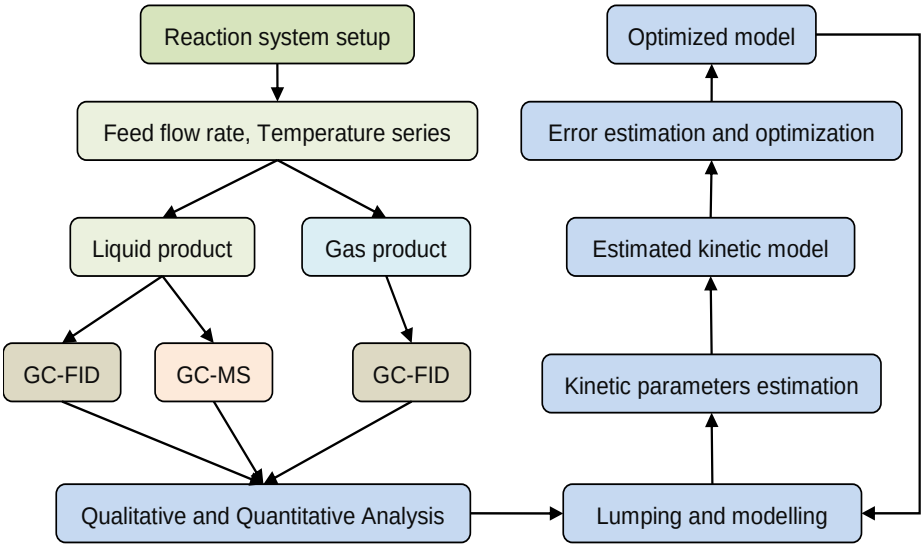


Figure 3-5. Flowchart of research method of the experiment procedure

At near atmospheric pressure, the reactions are designed and conducted under different feed flow rates and reaction temperatures. The reaction products were collected and analyzed by GC-FID-TCD and GC-MS as needed (Chapter 4). A lumped kinetic model was built and optimized based on the data gained from the qualitative and quantitative analysis of the product samples (Chapter 5). Multivariate data analysis is employed for further analysis of the reactions and optimization of the modeling (Chapter 6).

3.4.2 GC analysis of experiment samples

7890A GC-FID-TCD system and 7890A GC-MS system from Agilent Technologies Inc. are used for characterization of the liquid and gas samples at the reactor outlet. The feed gas flow is controlled by gas flow controllers from Brooks Automation and an AMD 1000 flowmeter from Agilent Technologies

was used to measure the gas product flow after cooling at the gas flow outlet in the experiments. All the tubing and fitting are ordered from Swagelok Company.

Details of setup of the GC-FID-TCD analysis for gas samples are shown in Table 3-4 and Table 3-5 and the detailed setup of the GC-FID-TCD and GC-MS for liquid samples are shown in Table 3-6 and Table 3-7.

Table 3-4. Setup of the GC-FID-TCD for the gas samples

Injection volume	≈ 1mL	
Purged packed inlet	Heater temperature	200℃
	Total gas flow	28ml/min
	Septum purge flow	3ml/min
	Temperature procedure	See Table 3-5.
Column	Size	10m×200μm×0.5μm
	Type	Agilent PLOT Alumina
FID	Heater temperature	250℃
	H_2 flow	60ml/min
	Air flow	400ml/min
	detector	Helium
	Post run	25ml/min

Table 3-5. Setup of the GC-FID-TCD for the gas samples

Steps	Start/ ℃	End/ ℃	Temperature increasing rate (℃/min)	Running time/min
1	70	70	0	7
2	70	250	10	18
3	250	250	0	2
4	250	70	-30	6
5	70	70	0	8

Table 3-6. Setup of the GC-FID-TCD and GC-MS analysis for the liquid samples

Injection volume	0.2 μL	
Purged packed inlet of the column	heater temperature	250 ℃
	Total gas flow	25ml/min
	Septum purge flow	2ml/min
	Temperature procedure	See Table 3-7.
Column	Size	50m ×200 μm ×0.5 μm
	type	Agilent 19091S-001
FID	Heater temperature	300 ℃
	H_2 flow	30ml/min
	Air flow	400ml/min
	detector	Helium
	Makeup Nitrogen	25ml/min

Table 3-7. Temperature Procedure of the GC-FID analysis for liquid samples at the injection inlet of the GC column

Steps	Start/ ℃	End/ ℃	Temp Increasing rate (℃/min)	Running time/min
1	40	150	5	22
2	150	150	0	5

Both the gas and the liquid samples condensed at the outlet of the reactor from the experiments were characterized by the GC-FID-TCD. GC-MS was also employed for characterization of the liquid samples and for identifying the degree of saturation of the chemicals at different chromatograph peaks of the liquid samples.

Further quantitative analysis was conducted with the GC-FID-TCD data with previous calibration. The method of the quantitative analysis of the mole ratios of the chemicals in the product samples are described as below:

1) Assume there are n peaks present in the GC chromatograph;
2) Collect all the chromatograph peaks from the GC experiments and their areas labeled as A_1, A_2, …, A_n;
3) For the peaks, the response factors are labeled as R_1, R_2,, …, R_n, from the calibration or from the reference;
4) The mole ratio $MR_i$of the chemical specie i:

$$MR_i = \frac{A_i \cdot R_i}{\sum_{j=1}^{n} A_j \cdot R_j}$$

The response factors for the calculation of gas samples are listed in **Table 3-8**. For the liquid samples, the value of the FID response factors for all the GC peaks is 1.00 for all and the peak areas are proportional to molar fraction, not mass fraction.

Table 3-8. Response factors for the calculation of GC of gas samples

Hydrocarbon Chemical	Response factor
methane	1.45×10^{-5}
ethylene	8.10×10^{-6}
ethane	8.25×10^{-6}
propylene	5.95×10^{-6}
propane	5.48×10^{-6}
butene	4.61×10^{-6}
butane	5.21×10^{-6}
pentane	1.56×10^{-5}
pentene	9.67×10^{-6}
hexene	2.16×10^{-5}
hexane	1.28×10^{-5}

When calculating the mole ratio of certain carbon number in the product, the mole ratio is the sum of carbon number in the gas sample and that in the liquid sample from the same experiment. For the gas samples, less peaks shown in the

chromatograph than the liquid samples where there are around 70 to 100 peaks. The GC-MS results are gained to assist in assignment of chromatograph peaks to different carbon numbers.

The mole ratio of certain carbon number ci is calculated as below:

$$\overline{MW} = \sum_{i=1}^{n} MR_i \cdot MW_i$$

$$MR_{ci} = \frac{\frac{Yrg}{\overline{MW_g}} \cdot MR_{ig} + \frac{Yrl}{\overline{MW_l}} \cdot MR_{il}}{\frac{Yrg}{\overline{MW_g}} + \frac{Yrl}{\overline{MW_l}}}$$

Where $\overline{MW}$ is average molecular weight of the gas sample of the liquid sample, MR_i is the mole ratio of chemical specie represented by peak i in the gas chromatograph, and MW_i is the molecular weight of this chemical specie, MR_{ci} is the mole ratio of carbon number ci in the whole reactant stream at the reactor outlet, Yrg is the yield rate of the gas by gram/min after condensation which is calculated from the volume flow rate by ideal gas law, $\overline{MW_g}$ is the average molecular weight of the gas sample, Yrl is the yield rate of the liquid by gram/min after condensation which is calculated by liquid yield divided by reaction time, MR_{ig} is the mole ratio of carbon number ci in the gas sample, $\overline{MW_l}$ is the average molecular weight of the liquid sample, MR_{il} is the mole ratio of carbon number ci in the liquid sample.

As the temperature T in unit of Kelvin of the gas at the inlet of flowmeter and the gas volume flow rate Vf in unit of ml/min were measured, the gas yield rate is calculated as:

$$Yrg = \frac{Vf \cdot 273.15}{T \cdot 22400} \cdot \overline{MW_g}$$

3.4.3 Programming and simulation

All the simulation and calculation are programmed and conducted by MATLAB. **Appendix E** and **Appendix F** give the detailed code for reference.

Chapter 4

Effect of space velocity and reaction temperature on the catalytic conversion of propylene over ZSM-5 zeolite

4.1 Introduction

This chapter discusses the effect of space velocity (weight hourly space velocity, WHSV) and reaction temperature on the catalytic conversion of propylene over ZSM-5 zeolite catalyst. This includes their effect on the carbon number distribution of hydrocarbons in the reactor outlet stream and the effect on the yields of the different lumps of chemical species including C_1-C_2 light gas which is an undesired by-product of the process, the desired C_{4+} hydrocarbons (including C_4 and excluding aromatics) and aromatics, the amount of which will highly affect the property of the liquid products condensed under room temperature.

4.2 Experimental procedure

The experimental setup and the method are as described in Chapter 3 of this study. Pure propylene and pure nitrogen from Praxair Technology Inc. are used as feed and purge gas respectively, for the experiments. Concentrations of the chemical species are measured by GC-FID-TCD and GC-MS from Agilent Technology.

0.15g ZSM-5 zeolite catalyst is packed inside the tubing reactor for the catalysis of the reaction. Before being blended with 400-mesh silicon carbide (SiC) for use in the reaction, the zeolite catalyst powder is activated by been calcined at 550 ℃ for 5 hours with dry air purging at 50ml/min. The blend ratios

of zeolites catalyst to silicon carbide is 1:20 by mass. During all the experiments, no catalyst deactivation was observed. No reaction was found in the experiments to occur without catalyst.

4.3 Temperature profile inside the reactor and pressure

The axial temperature profile was tested for all the experiments that were run, and with the small size of the reactor and short distance in the radial direction it is assumed that there is no radial temperature difference at the same axial location of the reactor. At temperatures higher than 300 ℃, this assumption works well and its effect on the reactant flow profile is insignificant theoretically and can be ignored under the reaction conditions. **Figure 4-1** shows three examples of the temperature profile inside the tubing reactor along its axis during the reactions in steady state. As is seen, the temperature in the catalytic reaction zone is relatively uniform, with a difference less than 3 ℃ from the beginning to the end of the catalyst zone insider the reactor. Thus to simplify the calculation, the reaction temperature is represented by the average temperature in the catalytic reaction zone in the later discussions.

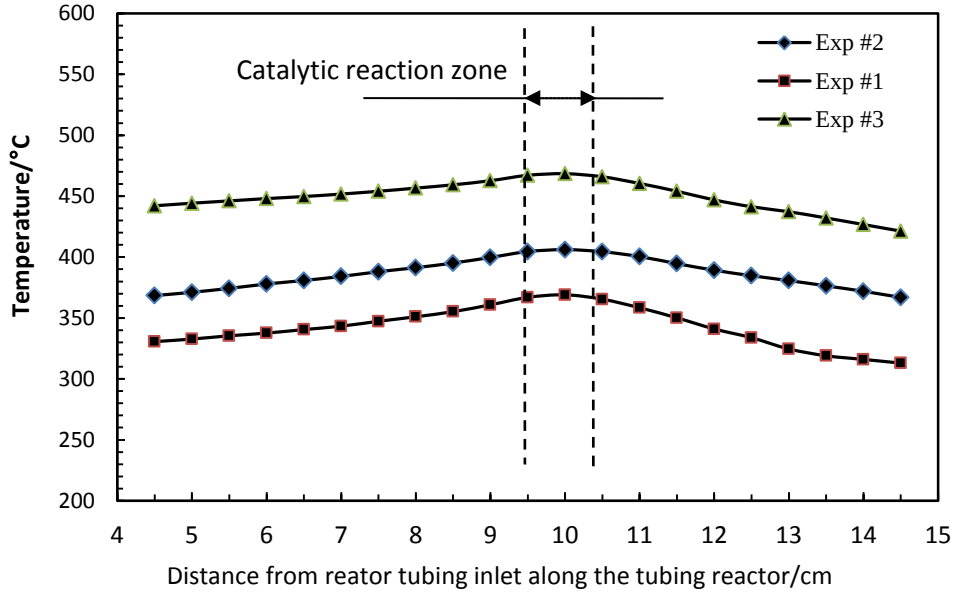


Figure 4-1. Temperature profile inside the reactor at steady state

All the experiments are carried out at near-atmospheric pressure inside the reactor. The only pressure profile is due to the pressure drop with difference less than 0.05atm, which is needed as driving force for enabling the continuous flow of the reactant stream.

4.4 Carbon number distribution

The carbon number distribution of the hydrocarbons at the reactor outlet is studies at steady state under different operating conditions. **Figure 4-2** shows the carbon number distribution of hydrocarbons at the reactor outlet based on molar fractions at different combinations of space velocity (WHSV) and reaction temperature at near-atmospheric pressure.

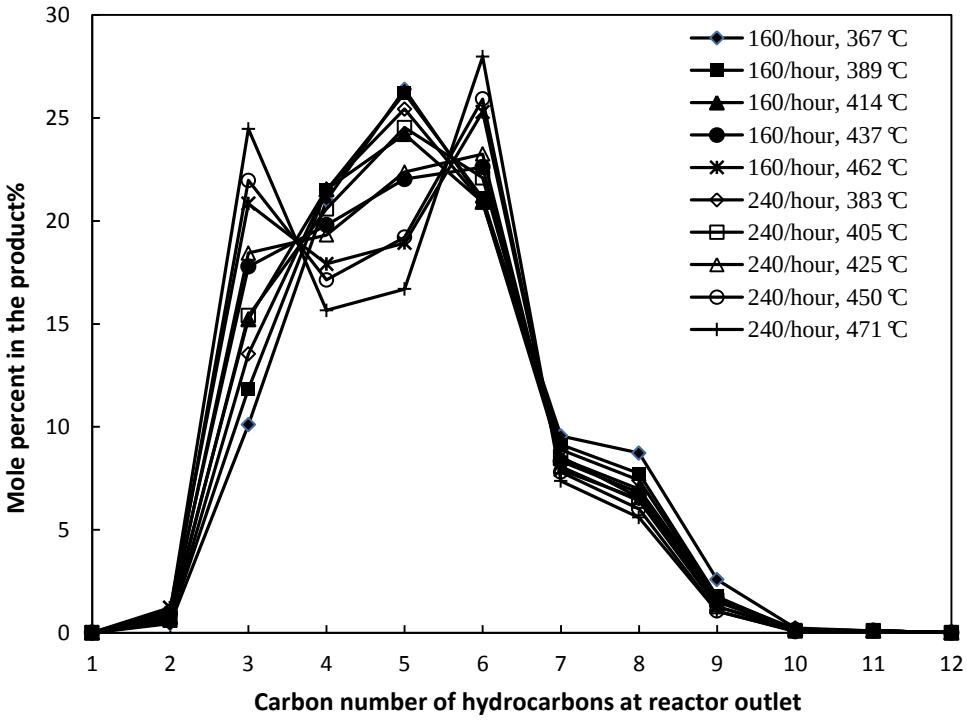


Figure 4-2. Carbon number distribution of hydrocarbons in the chemicals stream at the reactor outlet at different reaction temperatures and WHSV

It shows in **Figure 4-2** that the main products from the catalytic conversion of propylene under the experimental conditions are C_4 to C_9 hydrocarbons, which cover 70% to 90% of the product by molar fraction, with less than 2% of

C_1-C_2 light gas and C_{10}-C_{11} hydrocarbons also being present in the product. No C_{13} or heavier hydrocarbons are found in the hydrocarbon stream at the reactor outlet. This result agrees with the results of previous work reported in the literature on catalytic conversion of propylene over ZSM-5 zeolite catalyst at atmospheric pressure (Garwood, 1983). The mass recovery of propylene feed from the reactor outlet stream is in the range of 99% to 100% for all of the conducted experiments, which indicates that the mass loss during the reaction is in the range of allowable error. This implies that this process can be potentially used for conversion of propylene to gasoline range hydrocarbons with high material efficiency.

A list of reactions is proposed for this process as shown in **Table 4-1** according to the data of carbon number distribution above, where the reactions are classified as principal, secondary and minor according to the amount of products produced in the corresponding reaction. According to the results shown in **Figure 4-2**, the carbon is mainly distributed between C_3 and C_6, which results from the two principal reactions, #1 and #2, in **Table 4-1**. Further, through reaction #3 and #4, the carbon is redistributed among C_4 and C_5, leading to more C_5 being produced in the products than C_4. Meanwhile, the carbon distributed in C_7 to C_9 is produced from the secondary reactions #5 to #7 in **Table 4-**1 and reactions #9 to #13 explain the minor amount of C_1, C_2 and C_{10} to C_{12} existing in the conversion product as shown in **Figure 4-2**. Since there is a large difference in the yield of C_5 and C_7, C_4 and C_8, it is rarely possible for distribution of carbon through intermediate cracking reaction of C_{12} hydrocarbons which tends to produce equal amount C_5 and C_7, C_4 and C_8. Thus cracking of C_{12} hydrocarbons is considered as minor reaction path rather than the major devotion to the carbon number distribution.

Table 4-1. Important reactions in the process of propylene conversion

Label	Level of reaction	Reaction Formula
#1	Principal	$C_3+C_3 \rightarrow C_6$
#2	Principal	$C_3+C_6 \rightarrow C_4+C_5$
#3	Secondary	$C_4+C_4 \rightarrow C_3+C_5$
#4	Secondary	$C_3+C_5 \rightarrow C_4+C_4$
#5	Secondary	$C_3+C_4 \rightarrow C_7$
#6	Secondary	$C_4+C_4 \rightarrow C_8$
#7	Secondary	$C_3+C_5 \rightarrow C_8$
#8	Secondary	$C_4+C_5 \rightarrow C_9$
#9	Minor	$C_{10-n} + C_n \rightarrow C_{10}$ (n=3,4,5)
#10	Minor	$C_{11-n} + C_n \rightarrow C_{11}$ (n=3,4,5,6)
#11	Minor	$C_{12-n} + C_n \rightarrow C_{12}$ (n=3,4,5,6)
#12	Minor	$C_m \rightarrow C_{m-2} + C_2$ (m=6~10)
#13	Minor	$C_m \rightarrow C_{m-1} + C_1$ (m=6~10)

This agrees well with the view of Tabak et al. (1983) that the reaction of light olefins over ZSM-5 catalyst was found to proceed by reaction to discrete oligomers, followed by self-polymerization, copolymerization, cracking and aromatization. As the whole process goes less completely under high temperature and high space velocity, less C_6 will be converted to other hydrocarbons, thus leading to more C_6 being present in the product, as shown in **Figure 4-2**.

Also the data indicates that, with the same WHSV of 160/hour or 240/hour, the reaction products more tend to contain more C_4 to C_6 hydrocarbons with decreasing reaction temperature, while increasing the temperature will increase the amount of C_3 and C_6 in the product stream. Increasing the temperature will increase the reaction rate of all the reactions and decreasing WHSV will increase the residence time of the reactants in the catalytic zone. Thus as reactions of cracking and addition of olefins happen in the process with the same amount of catalyst employed, combination of high reaction temperature and low space velocity will lead to an approaching towards equilibrium of the reactions, and hence the carbon number distribution in the products. At the same space velocity, higher temperature accelerates both the cracking reaction and addition of olefins

at the same time, and will decrease the residence time as shown in **Table 4-2**. Thus, higher reaction temperature and higher WHSV leads to less C_3 to be converted, which agrees well with the results showed in **Figure 4-2**.

Table 4-2. WHSV and residence time for the different operating conditions

Exp #	WHSV/hour	Reaction temperature/ ℃	Residence time/sec
1	160	367	0.23
2	160	389	0.22
3	160	414	0.22
4	160	437	0.21
5	160	462	0.20
6	240	383	0.15
7	240	405	0.15
8	240	425	0.14
9	240	450	0.14
10	240	471	0.13

* The ideal gas law has been employed for calculation of the feed volumetric flow-rate in the catalytic reaction zone.

4.5 C_{4+} hydrocarbons

Figure 4-3 shows the mole fraction of C_{4+} hydrocarbons in the reactor outlet stream from reaction at different combinations of reaction temperature and feed rate (which is proportional to WHSV, since the same amount of catalyst is used in all the experiments).

It shows that the yield of C_{4+} hydrocarbons decreases with an increase in the WHSV in the range of experimental conditions studies. The C_{4+} yield increases with an increase in the reaction temperature from 330 ℃ to 370 ℃, and then decreases with an increase in the reaction temperature from 370 ℃ to 450 ℃. This is because a higher WHSV or feed rate leads to a smaller residence time of the reactant in the catalytic reaction zone, which directly leads to less conversion of propylene to heavier hydrocarbons. Also the change in the yield of C_{4+} hydrocarbons with the change in the reaction temperature is due to an increase in the reaction rate and a decrease in the residence time; the former is beneficial in increasing the yield of C_{4+} hydrocarbons, and the latter decreases the yield of C_{4+} hydrocarbons. In the temperature range from 330 ℃ to 370 ℃, the effect of

increasing reaction rate dominates over the effect from decreasing of residence time. However, in the temperature range from 370 ℃ to 450 ℃, the effect from decreasing of residence time dominates, thus decreasing the C_{4+} hydrocarbons yield. It is also shown in the figure that the yield of the C_{4+} hydrocarbons is more sensitive to the change in reaction temperature rather than in the change of WHSV. Increasing the reaction temperature will increase the volumetric flow rate according to the ideal gas law which leads to smaller residence time in the catalytic reaction zone of the reactants, thus further constraining the catalytic conversion of propylene.

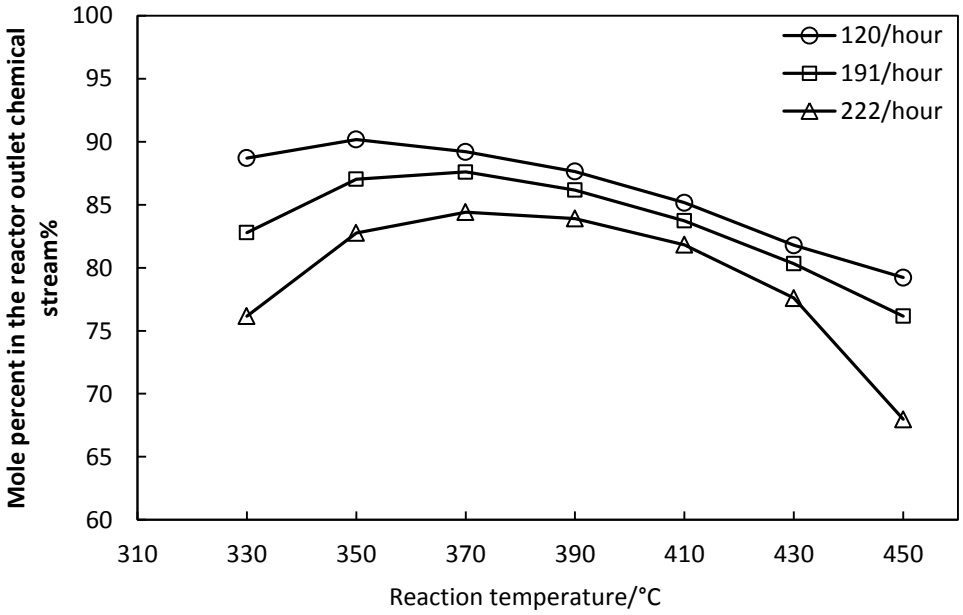


Figure 4-3. Mole fraction of C_{4+} hydrocarbons in the product stream at various WHSV and reaction temperature

4.6 C_1 and C_2 light gas yield

C_1 and C_2 hydrocarbons, including methane, ethane and ethylene, are the by-products of the catalytic conversion of propylene and are produced mainly from the cracking of heavy hydrocarbons. Their yields are of great interest since they are not only an important reference measure for evaluation of the process,

but are also valuable in estimating the appropriate temperature range for optimal operation of this process. As presented in **Figure 4-4**, the change in reaction temperature and space velocity will affect the yield of C_1-C_2 hydrocarbons. It is also implied from the figure that increasing the reaction temperature and decreasing the space velocity can both lead an increase in the C_1-C_2 yield. This is due to the fact that more cracking happens at higher reaction temperature; also, C_1-C_2 hydrocarbons are hard to convert to heavier hydrocarbons because of their stability, thus making the process almost irreversible towards the production of C_1 and C_2 at these higher reaction temperatures. The total yield of total C_1-C_2 also increases with a decrease in space velocity (WHSV) because of the smaller residence time, which reduces the catalytic conversion of propylene to heavier hydrocarbons followed by a lower rate of cracking of heavier hydrocarbons to C_1-C_2 light gas.

From **Table 4-3**, we also see that the main fraction of C_1-C_2 hydrocarbons is C_2. This is because cracking at α-carbon position generating C_1 is much harder than at β-carbon position (which produces C_2) due to the higher activation energy it needs. Increasing the reaction temperature will increase the ratio of C_1 in the C_1-C_2 hydrocarbons since it enables the higher activation energy pathway, but a smaller change in the C_2/C_1 ratio happens when increasing the WHSV as shown in **Table 4-3**.

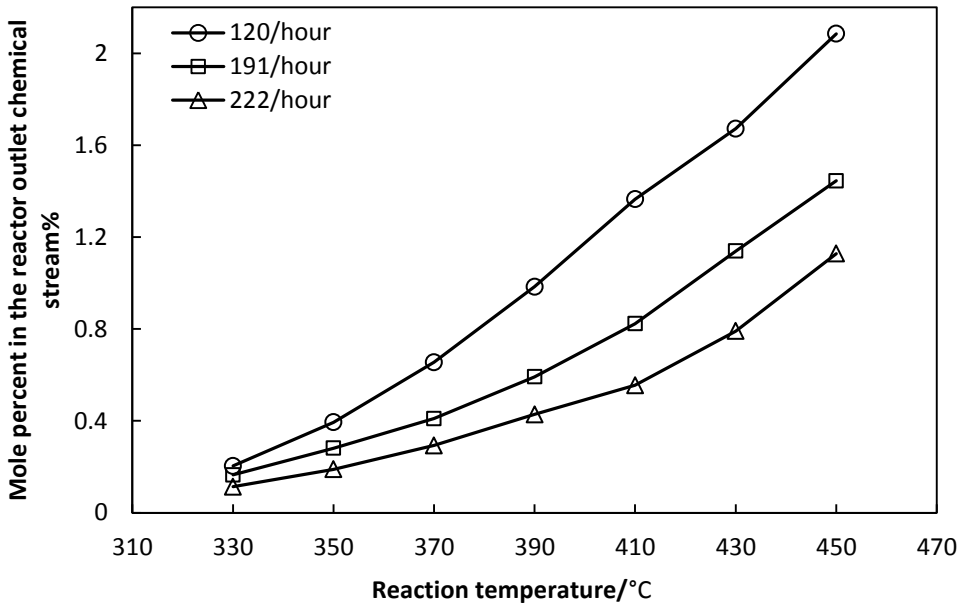


Figure 4-4. Mole fraction of C_1 and C_2 in the product stream at various conditions of WHSV and reaction temperature

Table 4-3. Mole fraction of C_1 and C_2 hydrocarbon in the product

Feed rate, mole/hour	Reaction temperature/ ℃	Mole fraction of C_1 in product%	Mole fraction of C_2 in product%	C_2/C_1
0.575	367	0.003	0.45	133
0.575	389	0.005	0.58	113
0.575	414	0.008	0.79	96
0.575	437	0.013	0.99	74
0.575	462	0.022	1.22	56
0.852	383	0.004	0.51	125
0.852	405	0.006	0.61	110
0.852	425	0.008	0.73	90
0.852	450	0.012	0.84	67
0.852	471	0.020	1.08	54

4.7 Aromatic yield

Aromatic yield is also an important reference for evaluation of the process of catalytic conversion of propylene to heavier hydrocarbons in the gasoline range. For the condensed liquid generated from catalytic conversion of

propylene, aromatic content will directly affect the property and quality of the condensed liquid product. So the yield of aromatics is of great interest and importance in study of this process. **Figure 4-5** shows the aromatic yield at different conditions of reaction temperature and space velocity.

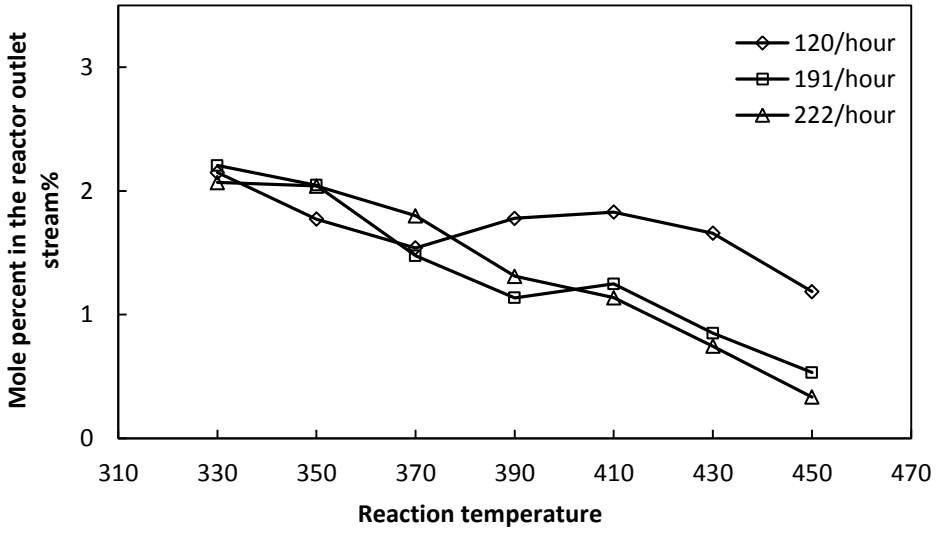


Figure 4-5. Mole fraction of aromatics in the reactor outlet stream different conditions of various reaction temperature and WHSV

Figure 4-5 indicates that at WHSV of 191/h and 222/h, aromatic yield first decreases with an increase in temperature up to 390 ℃, and then increases with a further increase in temperature up to 410 ℃, and then decreases again as the reaction temperature is increased until 450 ℃. A similar result at WHSV of 120/h, but the first change in the aromatic yield from decreasing to increasing with an increase in temperature occurs at 370 ℃ rather than at 390 ℃.

Also **Figure 4-5** suggests that under the same temperature, there is not much difference in aromatic yield at different WHSV or feed rate, which indicates aromatic yield of the process is not sensitive to the change of WHSV or feed rates, but is more sensitive to the temperature.

4.8 Conclusions

The effect of the temperature and feed rate on the conversion of propylene over H-ZSM-5 zeolite catalyst is discussed in this chapter.

A reaction mechanism is proposed for the process which explains the distribution of carbon numbers in the product. The experimental results indicate that, increasing of the reaction temperature will affect the carbon number distribution, leading to less conversion of propylene to heavier hydrocarbons and increase the yield of C_1-C_2 hydrocarbons under the same WHSV due to more cracking. The aromatic yield will decrease as the temperature is increased from 330 ℃ to 370 ℃ (low WHSV) or 390 ℃ (high WHSV); it then increases as the temperature is increased from 390 ℃ to 410 ℃ and it again decreases as the temperature is raised to 450 ℃. Additionally, the aromatic yield is much less sensitive to the WHSV than to the reaction temperature. Also, increasing of WHSV will decrease the conversion of propylene and the yield of C_{4+} hydrocarbons, C_1-C_2 and aromatics.

The low yield of C_1-C_2 hydrocarbons and aromatics each of which is less than 2.5% by mole fraction indicates this process has a high potential to be used for conversion of propylene to gasoline range hydrocarbons with high material efficiency.

The reaction mechanism that has been proposed here will be developed further in the next chapter, which focuses on lumped kinetic modelling.

References:

S. A. Tabak, F. J. Krabeck, W. E. Garwood. Conversion of propylene and butylene over ZSM-5 catalyst. AIChE Journal, 1983, 32:1526-1531.

W.E. Garwood. Conversion of C2 to C10 olefins into higher olefins on the synthetic zeolites ZSM-5. Erdol & Kohle Erdgas Petrochemie, 1982, 10(35): 479-497.

Chapter 5

Lumped kinetic modelling of the catalytic conversion of propylene over ZSM-5 zeolites catalyst

5.1 Introduction

Kinetic modelling is important not only for understanding the process of catalytic conversion of propylene, but is also practical for understanding the effect of reaction conditions such as reaction temperature and feed rate on the process, which further can be used to optimize the reaction process.

This chapter is devoted to building a lumped kinetic model of catalytic conversion of propylene over ZSM-5 catalyst, including its assumptions, the mathematics and the evaluation of the modelling results.

5.2 Reaction conditions and modelling assumptions

The experimental procedure is the same as described in Chapter 4 of this study. 0.15g ZSM-5 zeolite catalyst calcified at 550 ℃ for 5 hours in air blended with 3.00g 400-mesh silicon carbide is packed in the reactor for the catalysis of the reaction. Pure propylene is used as feed for the process under atmospheric pressure. The product is compressed and collected and is tested by gas GC and liquid GC individually. The mole fractions of the chemical species are calculated from the analysis of the gas chromatographic data.

The axial temperature profile is tested for all the experiments, and it is assumed that there is no radial temperature variation. **Figure 4-1** in Chapter 4 of

this study shows three examples of the temperature profile along the reactor during the reaction in steady state. It shows that the temperature in the catalytic reaction zone is relatively uniform (with variation less than 3℃); thus, to simplify the calculations, the reaction temperature is assumed to be the average temperature in the reaction zone in the later discussions on modelling.

Also, as discussed in Chapter 4, no hydrocarbons heavier than C_{12} are detected in the product, which supports first assumption that no deactivation of the catalyst happens in the experiments. Deactivation of the catalyst is caused primarily by the deposition of coke on the active sites of the catalyst. Thermogravimetric analysis (TGA) on the fresh catalyst and the used catalyst collected after reaction has been conducted to evaluate hydrocarbon deposition; the results are shown in **Figure 5-1** shows.

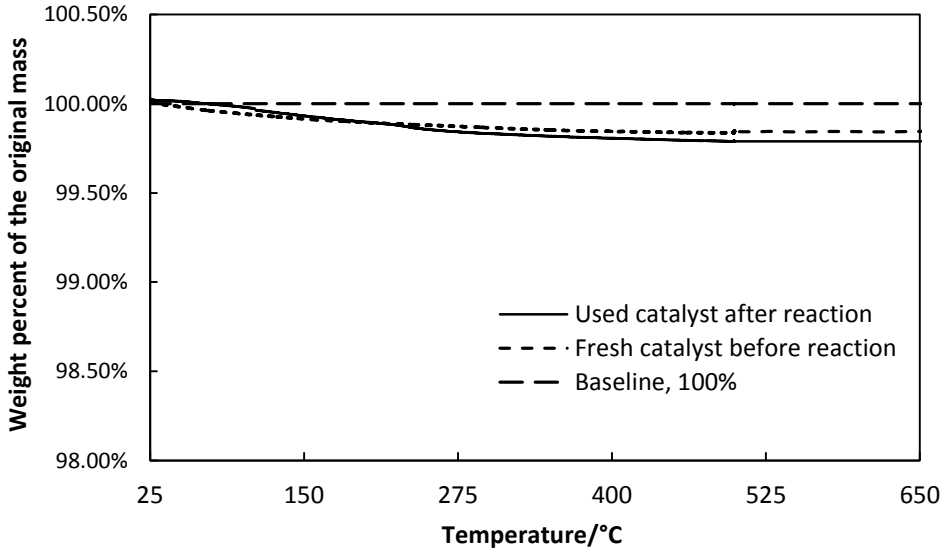


Figure 5-1. TGA analysis of fresh and used catalyst, performed between temperatures of 25℃ and 650℃, with the rate of temperature increase being at 10℃ /min and an air purge at 100ml/min

The TGA result of used catalyst after reaction and the fresh catalyst before reaction shows that there is almost no difference in the weight loss of these two catalyst samples and both have around 0.15% ~ 0.20% weight loss when being heated from room temperature up to 400℃ and then stay stable up to 650℃.

This indicates there is almost no deposit of hydrocarbons over the used catalyst and supports the assumption of no deactivation of the catalyst. The second assumption is that, there is no radial temperature profile and the reactant flow is ideally mixed isothermal plug flow. Also it is assumed that the reaction progress is constrained only by chemical reactions and the internal and external diffusions are neglected. The last assumption is the catalysts are uniformly distributed in the catalytic reaction zoom with ideally distributed active sites.

5.3 Lumping of chemical species in kinetic models

As mentioned in Chapter 1 of this study, a lumped kinetic study is necessary to simplify the reaction mechanism and make it practical to study the main reactions in the process. Theoretically, a comprehensive kinetic model should be constructed based on the reactions proposed in **Table 4-1**. However, kinetics based on that reaction mechanism do not explain the observed data well; also, the mechanism does not have the capacity to explain differences in aromatic yield, etc. Hence, we propose a different, more comprehensive lumped model based on six lumps in this study, the details of which are shown in **Figure 5-2**.

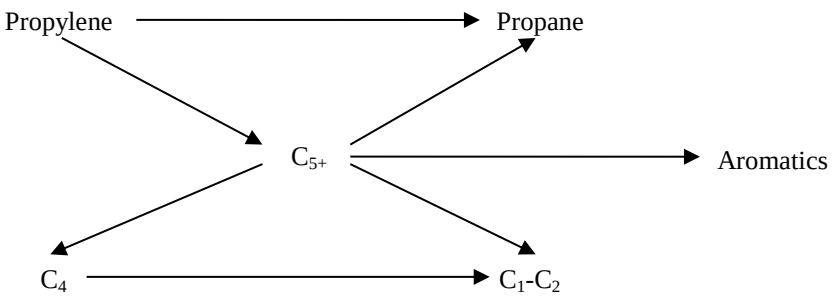


Figure 5-2. The complete 6-lump kinetic model network sketch

The six lumps that the chemical species involved are separated into are: propylene, propane, C_4 hydrocarbons, C_{5+} hydrocarbons (exclude aromatics), C_1-C_2 hydrocarbons and aromatics.

Propylene is the sole feed, so it is separated into a single lump. Propane is also placed in a single lump since the propylene feed can hydrogenate to form propane; also, cracking of heavier hydrocarbons can generate propane. Thus, the amount of propane is a good measure to evaluate the dehydrogenation process in the reaction network. C_1-C_2 hydrocarbons are lumped together; they may be generated from cracking reaction and the cracking property of the heavier hydrocarbons. C_4 hydrocarbons are lumped together based on our interest in tracking their yield, because of their ability to act as an additive in gasoline blending due to their. It is also reasonable to lump C_{5+} hydrocarbons into a single lump from the viewpoint of applications, as they are the main desired product and the major component of the liquid product. Thus, their yield can be used as an important measure of the potential liquid gasoline yield of the process. Aromatics are excluded from the C_{5+} hydrocarbons lump and treated as a separate entity, since their presence and mole fraction in the liquid will greatly affect the quality of the compressed liquid such as octane number, etc.

5.4 Modelling equations

With the assumptions above, the modelling equations consist of balances over the various species or lumps considered in the model.

Each lump reaction that converts material from one lump to another can be expressed as:

$$I \xrightarrow{k_{ij}} v_{ij} J \ , v_{ij} = \frac{\overline{MW_I}}{\overline{MW_J}} \text{ ---------- (1)}$$

where *I* and *J* are the two lumps involved in the reaction, k_{ij} is the reaction rate constant, v_{ij} is the stoichiometric coefficient of the reaction, and $\overline{MW_I}$ and $\overline{MW_J}$ is the average molecular weight of the lump *I* and *J*.

A materials balance over the feed rate can be expressed in the form of equation (2), with m_{acc} representing the accumulation of mass, m_{in} the inflow of mass, m_{out} the mass out flow, m_{gen} the mass generated and m_{con} the mass consumption.

$$m_{acc} = m_{in} - m_{out} + m_{gen} - m_{con} \text{ ---------- (2)}$$

At steady state, the mass accumulation is zero; thus, equation (2) becomes:

$$0 = c_i\, uAdt - (c_i\ + dc_i\,)(u + du)Adt + r_i Adxdt$$ ---------- (3)

where C_i is the concentration of lump *I* for any lump in the lumped kinetic model, u is the linear velocity of the reaction flow, A is the cross area of the reactor, r_i is the whole reaction rate of lump *I* when it is involved in more than one reactions, dx is the length difference and dt is the time difference.

After further simplification, we get:

$$udc_i\ + c_i\, du + dc_i\, du = r_i dx$$ ---------- (4)

The differential equation form of equation (4) is:

$$u \cdot \frac{dc_i}{dx} = \ r_i - c_i\ \cdot \frac{du}{dx}$$ ---------- (5)

Using the ideal gas law:

$$PV_t = n_t RT$$ ---------- (6)

where *P* is the pressure of gas flow, V_t is the gas volumetric flow rate, *R* is the universal gas constant, T is the temperature and n_t is the molar flow rate.

The gas velocity is expressed as:

$$\frac{du}{dx} = \frac{RT}{PA} \cdot \frac{dn_t}{dx}$$ ---------- (7)

Also we have that the change of concentration of any chemical lump is determined by the sum reaction rate of that lump in the axial direction:

$$\frac{dn_t}{dx} = \ \sum_1^i r_i \cdot A$$ ---------- (8)

Substituting equation (8) into equation (7) gives:

$$\frac{du}{dx} = \frac{RT}{P} \cdot \sum_1^i r_i$$ ---------- (9)

Substituting equation (9) into equation (5), then gives:

$$\frac{dc_i}{dx} = \frac{1}{u} \cdot (r_i - c_i\ \cdot \frac{RT}{P} \cdot \sum_1^n r_i)$$ ---------- (10)

For each reaction, the reaction rate constant is expressed according to Arrhenius equation:

$$k_{ij} = A_0 \cdot \exp\left({}^{E_a}/_{RT}\right)$$ ---------- (11)

where A_0 is the frequency factor, and E_a is the activation energy.

For each lump *i*, the total reaction rate is the sum of the rates of all the reaction involved:

$$r_i = \sum k_{ij} C_i^{m_{ij}} \quad \text{---------- (12)}$$

Equations (10), (11) and (12) are the basics for the kinetic modelling. The concentrations of lumps are obtained from gas chromatography (GC) data. For each reaction, there are three undetermined kinetic parameters, the frequency factor, the activation energy and the reaction order. Thus there are 15 kinetic parameters to be estimated with the experimental data. 25 sets of data are used for the estimation. The objective function to be minimized is:

$$f_{min} = \sum_{i=1}^{i=m} \sum_{j=1}^{j=n} \left(C_{ijcal} - C_{ijexp}\right)^2 \quad \text{---------- (13)}$$

where m is the number of experiments, n is number of lumps, C_{ijcal} is the calculated concentration by the modelling and C_{ijexp} is the concentration from the GC test.

The minimization of the objective function was carried out using the function 'fmincon' in Matlab. Certain constrains are placed on the calculation of the min value of objective function. The final results are shown in the next section.

Different reaction orders were hypothesized for each of the reactions, and the combination of reaction orders that had the best fit based on minimization of modelling error was chosen.

5.5 Modelling results

5.5.1 Kinetic parameter estimation

25 sets of experimental data are employed in the modelling for the estimation of the kinetic parameters. The estimated values of the kinetic parameters are shown in **Table 5-1**.

Table 5-1. Estimated value of parameters in the lumped kinetic model

Reaction	Reaction order	Frequency factor	Activation energy (KJ/mol)
Propylene → Propane	1.71	1.08×10^{4}	154.0
Propylene → C_{5+}	2	2.72	6.08
C_4 → C_1-C_2	1.20	4.45×10^{3}	109.7
C_{5+} → C_1-C_2	1	4.41×10^{2}	59.4
C_{5+} → propane	1	2.31	25.2
C_{5+} → C_4	2	6.92	27.3
C_{5+} → Aromatics	1	1.65	14.3

5.5.2 Modelling results

With the estimated kinetic parameters shown in **Table 5-1**, the predicted concentration of the lumps in the product and the concentration data from experiments are shown in **Figure 5-3** and **Figure 5-4**.

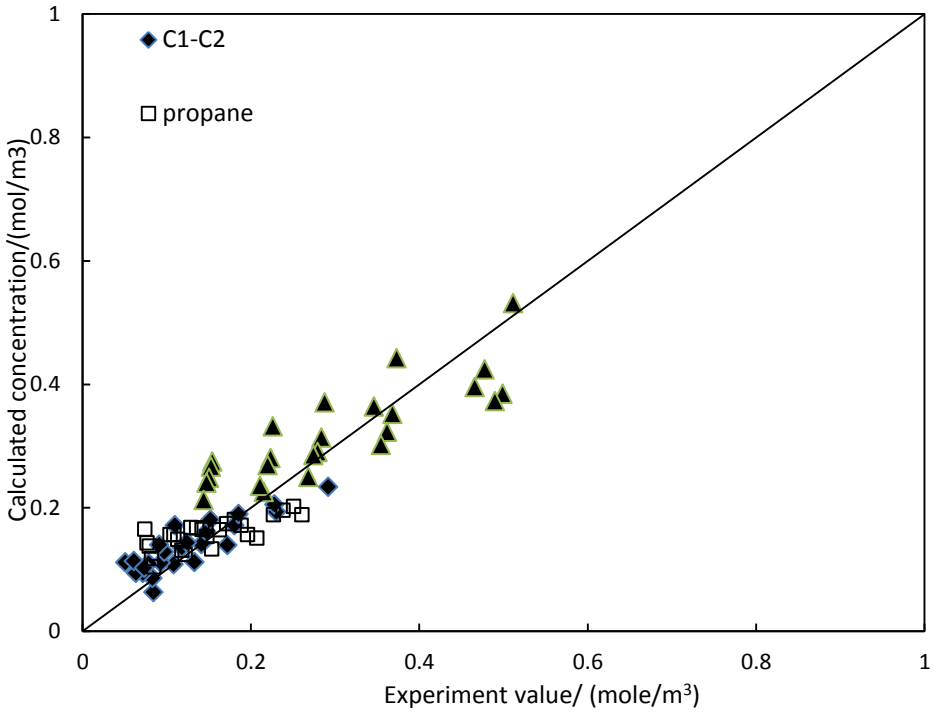


Figure 5-3. Modelling calculated value vs. experimental value of concentration of C_1-C_2, propane and aromatics in the product

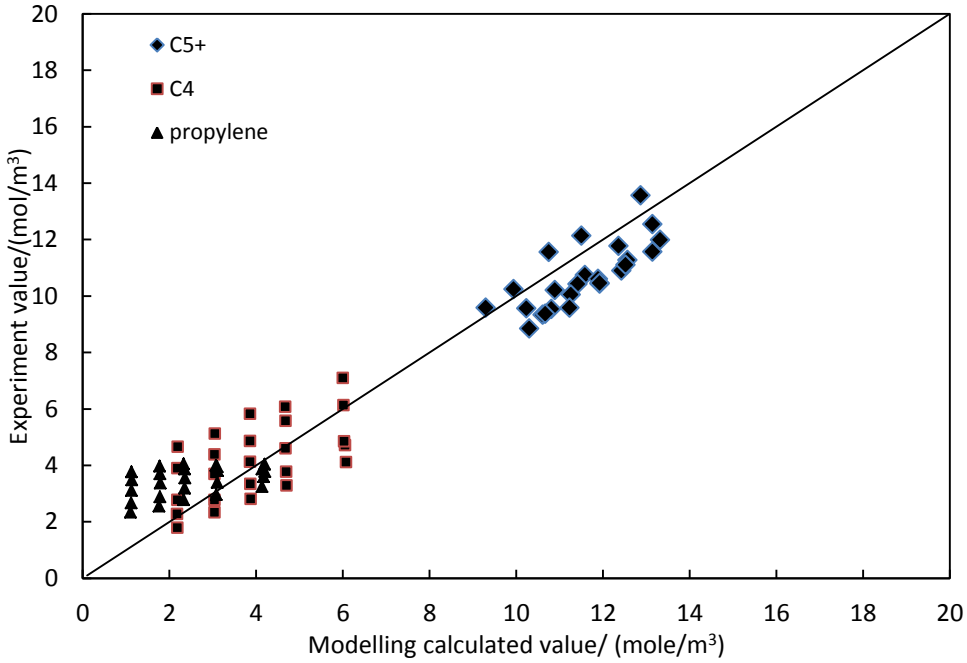


Figure 5-4. Experiment value VS modelling calculated value of the concentration of propylene, C_4 and C_{5+} hydrocarbons in the reactor outlet stream

From both **Figure 5-3** and **Figure 5-4**, they show that the built model can predict the concentration of the built lumps under different reaction conditions within allowed experimental error.

With the estimated kinetic parameters in **Table 5-1**, the concentration of the six lumps from experiments under different reaction temperature and space velocity are predicted and compared with the experimental data as shown in **Figure 5-5** and **Figure 5-6**. As shown in the figures, the prediction value by the lumped kinetic model agrees well with the experiment value with less than 15% error, mostly around 6% to 13%. It further proves that the built lumped kinetic model is a good reference for prediction of concentration of the lumps in the experiment under different reaction temperature and feed rate.

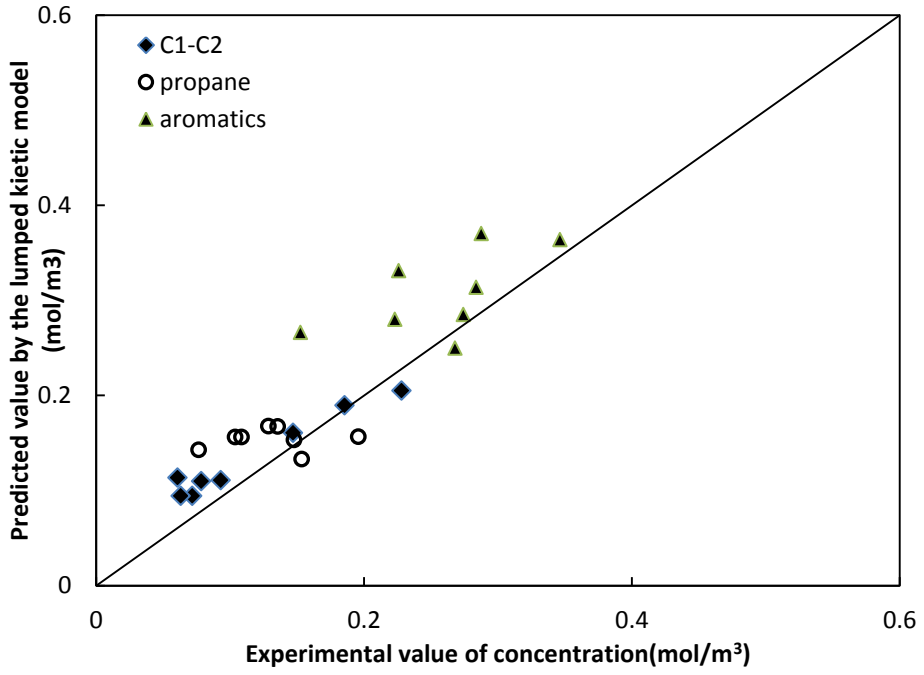


Figure 5-5. Concentration at the reactor outlet of C_1-C_2, propane and aromatics from experiments vs. the predicted value by lumped kinetic model

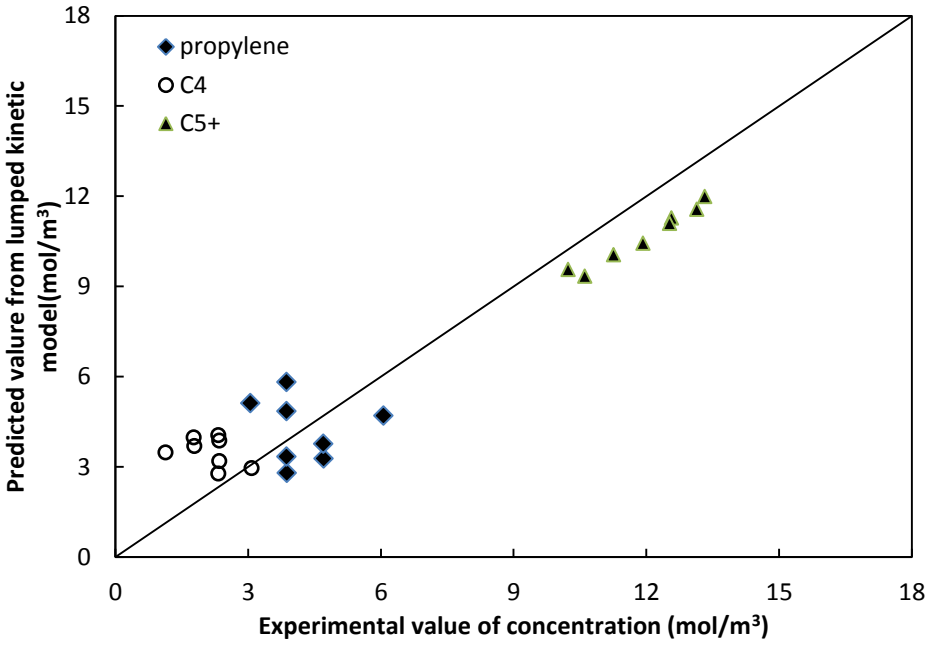


Figure 5-6. Concentration at the reactor outlet of propylene, C_4 hydrocarbons and C_{5+} hydrocarbons (excluding aromatics) from experiments vs. the predicted value by the lumped kinetic model

5.6 Conclusion and summary

A lumped kinetic model is built for the catalytic conversion of propylene over ZSM-5 zeolites catalyst. The undetermined kinetic parameters are estimated and the modelling result shows a good agreement with the experimental data. The prediction of concentration of the six chemical lumps by the built lumped model agrees well with the experimental value which further proves the high potential of the built model in practical use for experimental prediction.

Further discussion will be held in next chapter as it will discuss about combination of the modelling with the multivariate analysis for analysis and control of the process of catalytic conversion of propylene.

Chapter 6

Multivariate data analysis and spectral analysis for process control of the catalytic conversion of propylene over ZSM-5 zeolite catalyst

6.1 Introduction

Multivariate data analysis is useful in the extraction of latent (potentially existing but not presently realized) information from a high-dimensional data space and simplifying the analysis by reducing the number of relevant factors. In this chapter, principal component analysis (PCA) and hierarchical clustering analysis (HCA) (described in Chapter 2) are employed to analyze data to further understand the process of catalytic conversion of propylene over ZSM-5 catalyst. The analysis has potential to be used in the design of process control algorithms for the system.

Spectroscopic analysis, including Fourier transform infrared (FTIR), near infrared (NIR) and ultraviolet-visible (UV-Vis) spectroscopy, is a quick and informative method to analyze complex chemical reaction systems. Combined with multivariate data analysis, online spectral data can be used for real-time analysis and process control without requiring much calibration.

This chapter first devotes some attention to employing multivariate data analysis on the process of catalytic conversion of propylene over ZSM-5 catalyst, and then on combination with spectral data to provide guidelines for developing a process control strategy.

6.2 Data pre-processing

6.2.1 Data source

The data concerned in this chapter includes the offline gas chromatographic (GC) data of product samples from the experiments described in Chapter 4, and also offline and online spectral data on the liquid samples (including FTIR, NIR and UV-Vis).

The collected samples of IR, NIR and UV-Vis spectra showed that the IR spectra contain all the information with spectral band responses from alkane, alkene and aromatic hydrocarbons, but the NIR and UV-Vis do not show much information from the hydrocarbons and has a high noise level. This is mainly because the hardness of even flow designs of the condensed liquid through the flow cells which bring much noise and meaningless spectral data and also the capability with the equipments. Due to this, only the IR spectra data are used in further discussions.

6.2.2 Data pre-processing

Before applying multivariate data analysis, all the spectral variables are first rescaled and centered to eliminate the baseline drift and reduce the evenly distributed noise signal, and then normalized to the range of 0 to 1. Meanwhile, only the concerned IR bands listed in **Table 6-1** are extracted from the original spectra data are analyzed.

Table 6-1. The IR wave number bands of functional groups in hydrocarbons

Functional group	IR band (/cm)	Functional group	IR band(/cm)
Alkane		Alkyne	
-C-H stretch	3000~2950	C-H stretch	3300
-C-H bending	1480~1350	-C≡C- stretch	2260~2100
Alkene		Aromatic	
=C-H stretch	3100~3010	C-H stretch	3100~3000
=C-H bending	1000~675	C=C stretch	1600~1585
C=C stretch	1680~1620		1500~1400

6.3 Multivariate data analysis of carbon number distribution

25 sets of data with each including temperature, residence time calculated from the kinetic model as built in Chapter 5, feed molar flow rate and the concentrations of C_1 to C_{10+} chemical groups (as shown in Chapter 4 for the carbon number distribution) are analyzed by PCA. The temperature, feed molar rate are the manipulated inputs of the reaction system which are specified in the experiments. The residence time is determined by combination of these two inputs under the constraint of the built lumped kinetic model in Chapter 5 of this study. The conversion of propylene and the outlet concentrations of the chemical groups from C_1 to C_{10+} are all taken as the outputs of the reaction system.

14 principal components are gained with the first 4 principal components in total accounting for 98% of the whole variance as shown in Table 6-2 on analysis of reaction conditions of temperature, feed molar rate, residence time, and propylene conversion, concentrations of the C_1 to C_{10+} hydrocarbons. **Table 6-2** also shows the loadings for the first 4 principal components of the selected variables and **Figure 6-1** shows the loading for the first principal component (PC1) vs. against that for the second principal component (PC2) for each of the 10 carbon number hydrocarbon groups, the temperature and the residence time. The loading of a variable on a principal component indicates how the variable correlates with that component. A positive or negative loading on a principal component means the contribution from that variable on the principal component is in the positive or negative direction, respectively, and the absolute value of the loading indicates how much of that variable contributes to the variance of the principal component.

As shown in **Table 6-2**, the first principal component accounts for 73% of the total variance, with the loadings being evenly distributed over all the variables in the range of 0.20 to 0.30 (positive or negative). This means that the main variance mostly results from the direction of increasing temperature and feed molar rate, which have positive loadings 0.29 and 0.19, respectively. As the data indicates, the concentration of C_1, C_2, C_3 and C_6 in the reactor outlet chemical stream whose loadings are positive increases with increasing of

reaction temperature and feed mole rate, but in contrast does the propylene conversion, the residence time and concentrations of other carbon number groups whose loadings are negative. Increasing the residence time will increase the feed conversion in a system of irreversible reactions or one at conditions far away from equilibrium. In this case, the concentrations of C_1 and C_2 increases with an increase of reaction temperature and a decrease in the residence time, which theoretically lowers the conversion of propylene to C_1 and C_2. This suggests that the main factor for C_1 and C_2 production is temperature rather than the residence time. Concentration of C_3 and C_6 increases with an increase of in reaction temperature and feed flow rate due to the effect of decreasing the residence time, which reduces the extents of the reactions following C_6 formation from addition of two propylene molecules; this finally leads to less conversion of C_6. This also suggests that the reactions in the reaction network proceed sequentially with the first step of addition of C_3 to form C_6 followed by further addition reactions or cracking reaction to produce other carbon number chemicals, which agrees well with the discussion in Chapter 4.

Table 6-2. Results for Principal components analysis of carbon number distribution

Components	**1st PC**	**2nd PC**	**3rd PC**	**4th PC**
Of total variance	**73%**	**17%**	**7%**	**1%**
Accumulated variance	**73%**	**90%**	**97%**	**98%**
Temperature	**0.29**	-0.20	0.14	0.01
Feed molar rate	**0.19**	**0.48**	**-0.19**	**-0.33**
Residence time	**-0.22**	**-0.42**	0.04	**0.41**
$C_{3=}$ conversion	**-0.31**	0.03	0.08	0.17
C_1	**0.23**	**-0.43**	0.12	-0.25
C_2	**0.21**	**-0.43**	**0.33**	-0.24
C_3	**0.31**	-0.04	-0.07	-0.17
C_4	**-0.27**	0.10	**0.47**	-0.21
C_5	**-0.30**	0.15	0.19	0.10
C_6	**0.28**	-0.04	**-0.38**	**0.44**
C_7	**-0.29**	-0.02	-0.01	-0.36
C_8	**-0.29**	-0.17	-0.16	0.08
C_9	**-0.28**	-0.16	**-0.31**	-0.18
C_{10+}	**-0.21**	**-0.29**	**-0.53**	-0.37

The second principal component accounts for 17% of the total variance. Positive loading from feed flow rate and negative loading from temperature and negative loadings from residence time and concentrations of C_1, C_2 hydrocarbons indicate that increasing the feed flow rate and decreasing the temperature strongly correlate with decreasing the residence time and the concentration of C_1 and C_2 hydrocarbons. The third principal component which accounts for 7% of the whole variance indicates that decreasing the feed flow rate and increasing the temperature to some degree correlate with increasing the product concentrations of C_2 and C_4 and decreasing concentrations of C_6, C_9 and C_{10+}. Though accounting for only 1% of the whole variance, the fourth principal component implies that the effect of increasing the feed mole rate and residence time tends to increase the yield of C_6, and decreases the yield of C_1, C_2 and C_{10+}.

Beyond PCA, hierarchical clustering analysis was performed on the loadings of the first two principal components. This shows the similarity in the variance of the individual carbon number hydrocarbons at different reaction conditions (see **Figure 6-2**).

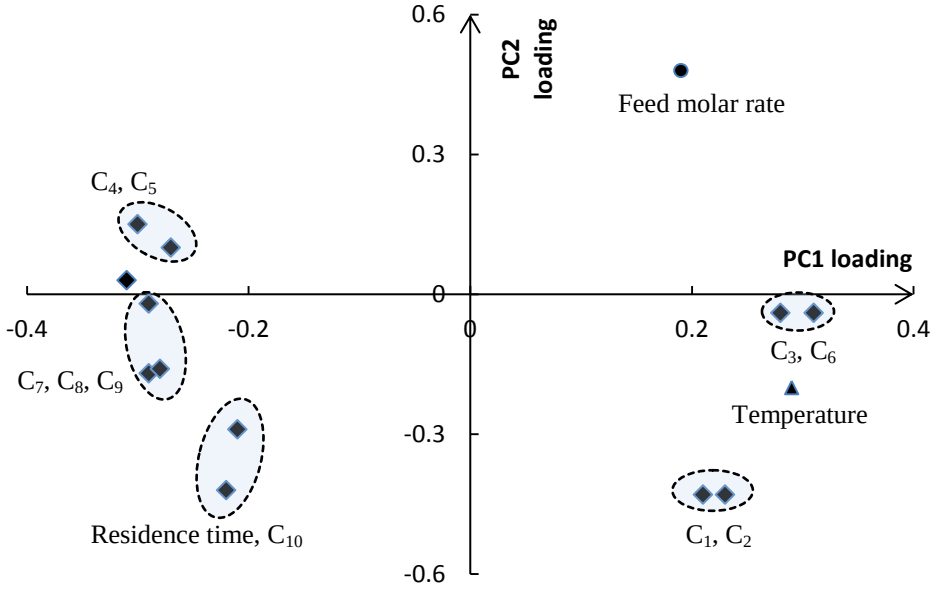


Figure 6-1. A loadings plot in terms of PC1 (x-axis) vs. PC2 (y-axis) for the factors considered in **Table 6-2**

HCA employs the distance calculation to evaluate the similarity of the objects and then groups them into clusters step by step with the objects in the same cluster are more similar than that in the others. More details were described in Chapter 2.

It shows in **Figure 6-2** that the lowest level clustering forms groups from pairs of C_1 and C_2, C_3 and C_6, C_8 and C_9. This indicates similar reactions or process could happen to these carbon number pairs, and thus offers valuable information for lumping algorithm when attempting to build improved lumped kinetic models from a statistical view. One example is choosing a cutting value of 1.5 in **Figure 6-2** and then classifying the carbon numbers into 5 lumps: lump 1 with C_1 and C_2, lump 2 with C_3 and C_6, lump 3 with C_4 and C_5, lump 4 with C_8 and C_9, and lump 5 including C_{10+} only. Future work needs to be done for consummation and validation of this idea for optimization the lumped kinetic modelling.

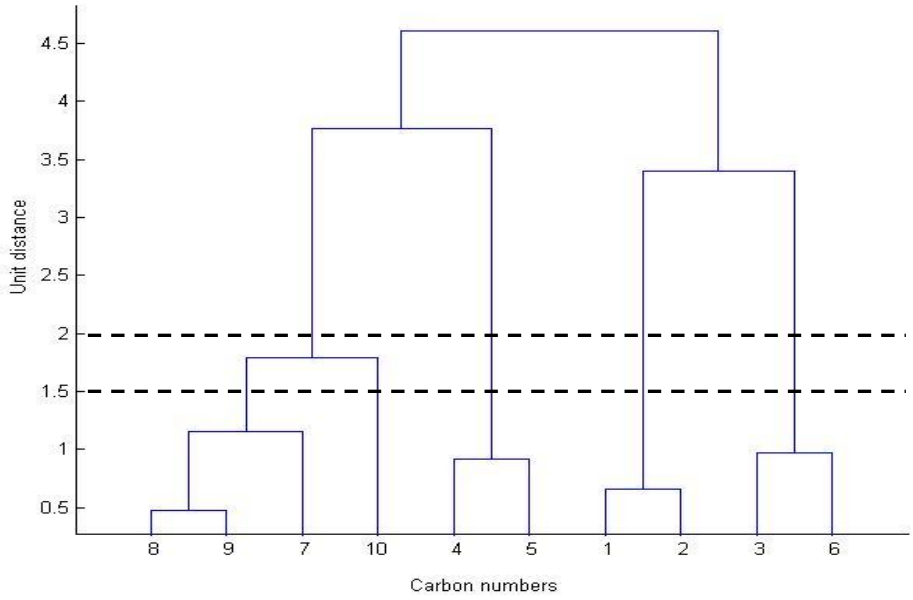


Figure 6-2. Hierarchical clustering analysis (HCA) of PC1 and PC2 loadings of the carbon number concentrations from C_1 to C_{10+}, the carbon number is labeled by its corresponding number in the figure and C_{10+} is represented by number 10 as shown.

6.4 Multivariate data analysis of the spectral data

6.4.1 Mapping of the spectra pattern to the reaction condition

The mapping of system parameters including the reaction conditions with the spectral patterns is critical in employing spectral data for analysis and control of the chemical process. As reaction condition is specified for a chemical reaction system, the system finally will be in a steady state with all state variables such as the reaction networks, the reaction rates and the reactant concentration profile and so on being constant, which further reflects on the product information such as product concentrations, product flow rate and etc. With spectral testing on the product produced under a specified reaction condition, a unique mapping can be built between the spectral pattern and the As **Figure 6-3** shows, the qualitative and quantitative information of the chemical samples from the reactions is reflected in the spectral pattern, which can be further projected to the system state and reaction conditions.

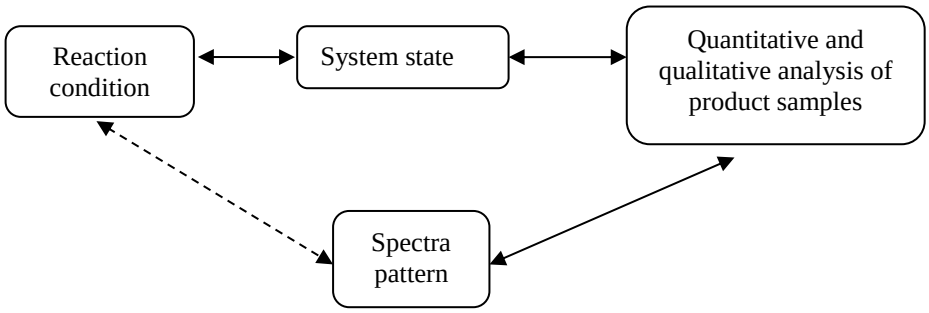


Figure 6-3. Mapping relations in the reaction system

25 sets of IR spectra data are analyzed by HCA and the dendrogram is shown in **Figure 6-5**. The IR patterns are shown in **Figure 6-4** and the samples are labeled under the corresponding reaction conditions listed in **Table 6-3**. Even no obvious similarity can be detected visually in **Figure 6-4**, it is shown in **Figure 6-5** and **Figure 6-6** that the clustering structure of the spectral data is such that similar reaction conditions are clustered together.

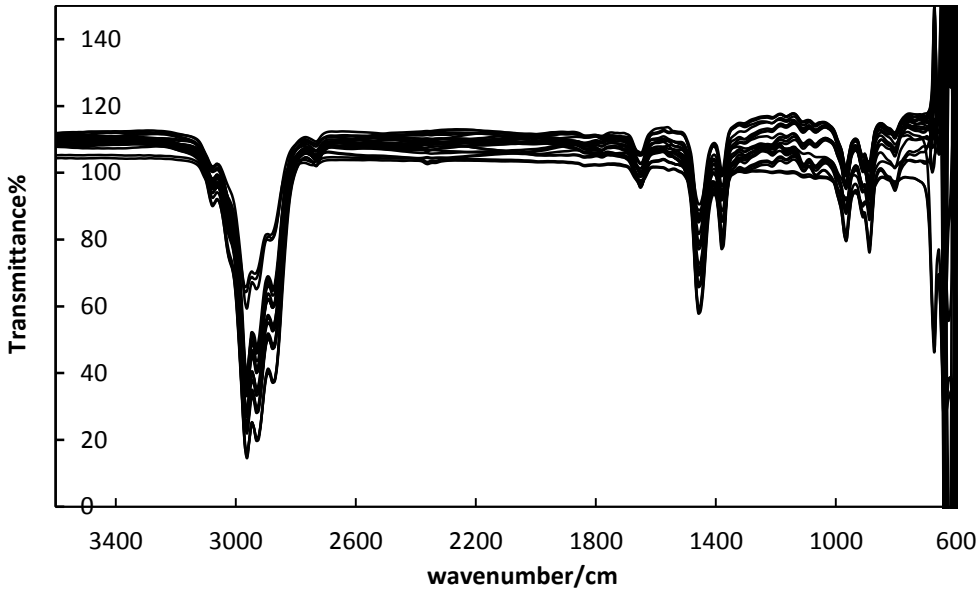


Figure 6-4. 25 sets of IR spectra of liquid product samples in the range 3500~600cm^{-1} at various reaction conditions of feed rate and reaction temperature as listed in **Table 6-3.**

Table 6-3. Experiment conditions of feed rate and reaction temperature

Exp #	Reaction Temp/ ℃	Feed MR Mole/min	Exp #	Reaction Temp/ ℃	Feed MR Mole/min
1	347	0.0071	14	426	0.0142
2	369	0.0096	15	435	0.0182
3	378	0.0119	16	434	0.0071
4	387	0.0142	17	439	0.0096
5	400	0.0182	18	444	0.0119
6	383	0.0071	19	451	0.0142
7	391	0.0096	20	458	0.0182
8	400	0.0119	21	459	0.0071
9	406	0.0142	22	464	0.0096
10	417	0.0182	23	469	0.0119
11	406	0.0071	24	473	0.0142
12	415	0.0096	25	479	0.0182
13	422	0.0119			

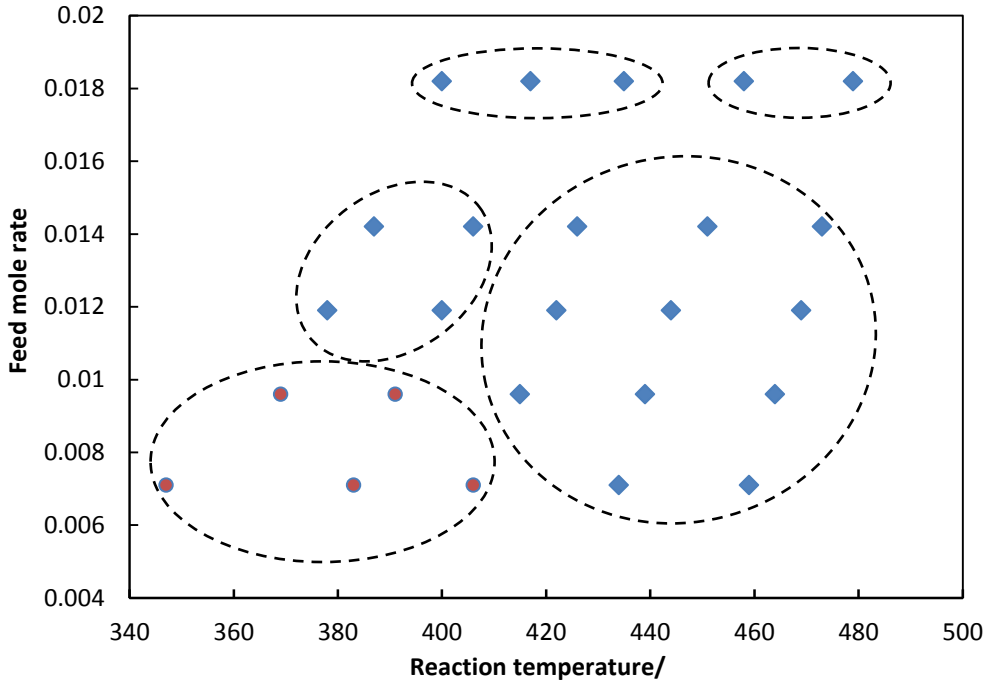


Figure 6-5. Plot showing temperature and feed molar rate, with the circles corresponding to the clusters in the 5-cluster structure in **Figure 6-6**

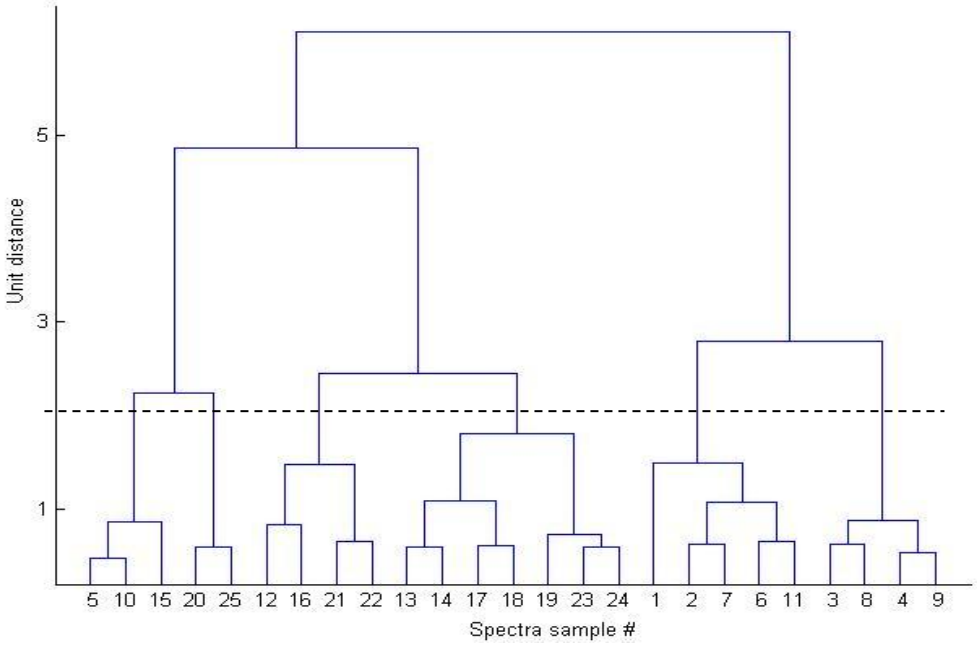


Figure 6-6. Dendrogram of HCA on the spectral data of liquid product from experiments at conditions listed in **Table 6-3**

To test the accuracy of this clustering structure with respect to reaction temperature, 12 new sets of IR patterns of liquid products in range of 3500~600cm^{-1} from experiments at conditions listed in **Table 6-4** are tested by HCA.

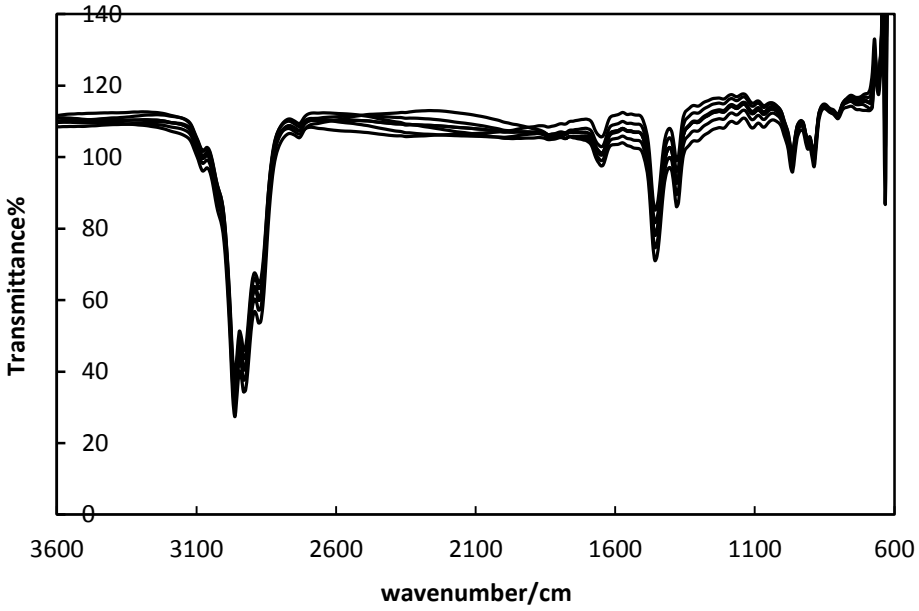


Figure 6-7. 12 New Spectral patterns of liquid products from validation date obtained at conditions listed in **Table 6-4**

To simulate system states of temperature vibration within 4 ℃ at set reaction temperatures, three separate experiments were conducted with all temperature increase of 4 ℃ per 30min and start temperatures of 393 ℃, 403 ℃ and 413 ℃ individually. In each experiment, IR spectral test of the liquid product was conducted every 6 minutes with the first test at any time of the first 5 minutes of the procedure and 4 tests taken for analysis. The labels of the spectral test results are as shown in **Table 6-4**.

Table 6-4. Experiment conditions of feed rate and reaction temperature

Exp #	Reaction Temp/℃	Feed MR Mole/min	Exp #	Reaction Temp/℃	Feed MR Mole/min
1	395±2	0.0119	7	405±2	0.0119
2	395±2	0.0119	8	405±2	0.0119
3	395±2	0.0119	9	415±2	0.0119
4	395±2	0.0119	10	415±2	0.0119
5	405±2	0.0119	11	415±2	0.0119
6	405±2	0.0119	12	415±2	0.0119

With baseline drift and variance in light source intensity and other noise sources, it is very difficult to distinguish between or observe patterns in of the IR patterns without statistical analysis. The clustering structure of the 12 sets of IR data is shown in **Fig 6-8**.

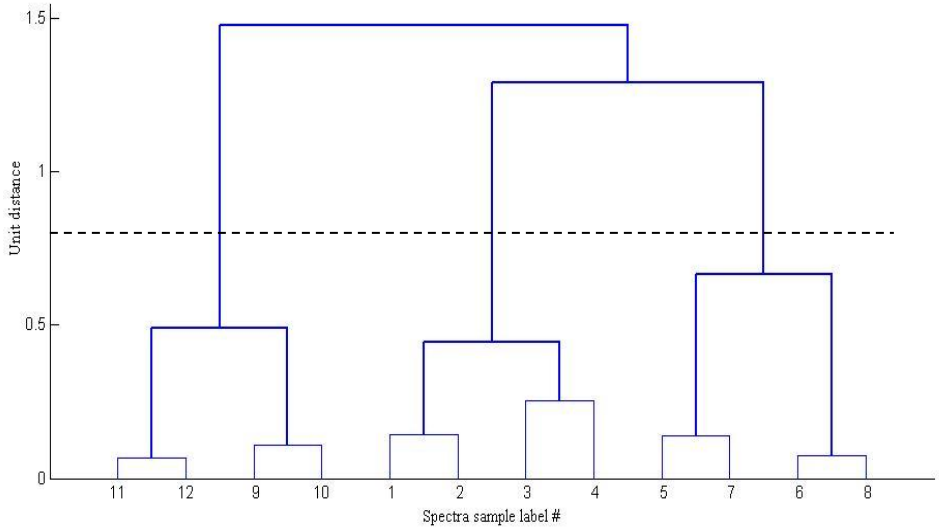


Figure 6-8. Clustering result of the 12 sets of spectral data liquid samples with the product for the validation data; the experimental conditions are listed in **Table 6-4**

It is seen in **Fig 6-8** that, the 12 sets of spectra patterns are clustered into 3 groups by any cutting value between 0.6 and 1.2 with each group containing exactly 4 samples as classified by the reaction temperature range shown in **Table**

6-4, where each cluster corresponds to each experiment as described above. This further proves that mapping of spectra pattern to the reaction condition works well by HCA, which forms the basis of discussion provided later in section 6.4.2. Due to the limitations in the temperature control of the furnace and other experimental factors, no further exploring in clustering structure in accuracy of reaction temperature has been conducted.

6.4.2 HCA of spectra data for outlier detection

On the basis of the mapping developed between spectral patterns and reaction conditions, an algorithm is proposed for outlier detection by HCA using online spectra and comparing against the clusters of the historical spectral data from normal operation of the system which is shown in **Figure 6-9**.

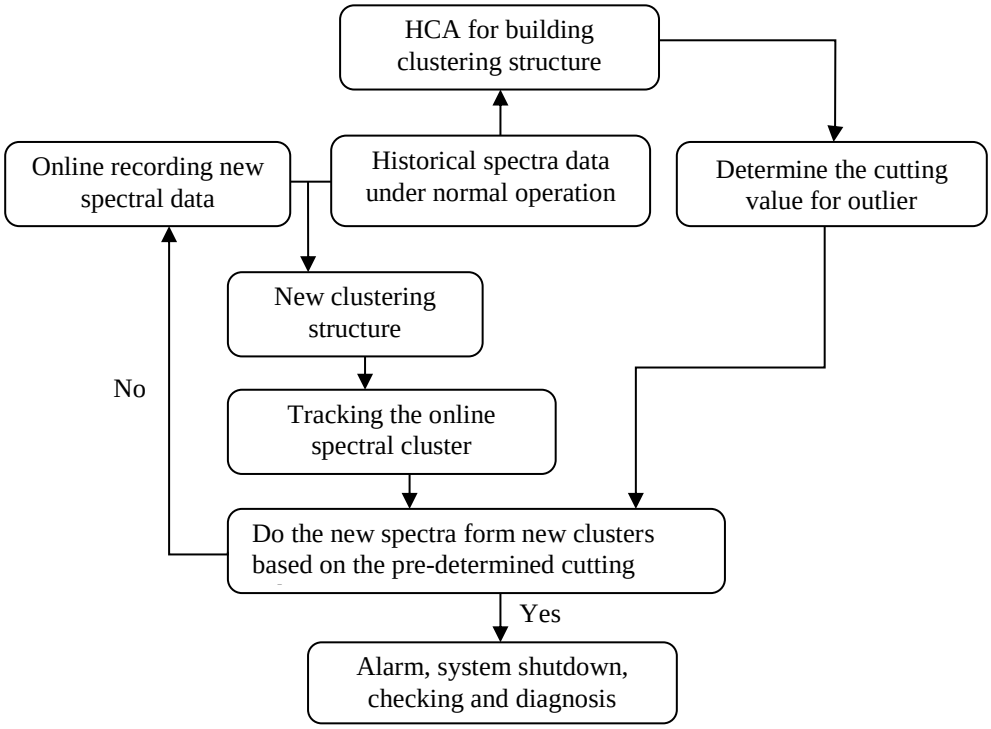


Figure 6-9. Algorithm for on-line outlier detection by HCA based on historical spectra data and online spectra data

The essence of the algorithm in **Figure 6-9** is to use the historical spectral patterns as standards for comparison with the online spectral data to determine if the system is currently in a normal operating condition. As new spectral patterns

are inserted into the cluster structure of the historical data under the same reaction conditions, any formation of a new cluster could mean deviation from the normal state beyond the allowable range; this can be judged as an outlier and an alarm can be turned on for system checking and diagnosis.

6.4.3 HCA of spectral data for process control

On the same basis of mapping of the spectra pattern and the reaction conditions such as temperature and feed flow rate, a process control algorithm is also proposed for the reaction system as presented in **Figure 6-10**, where reaction temperature is taken as the controlled variable.

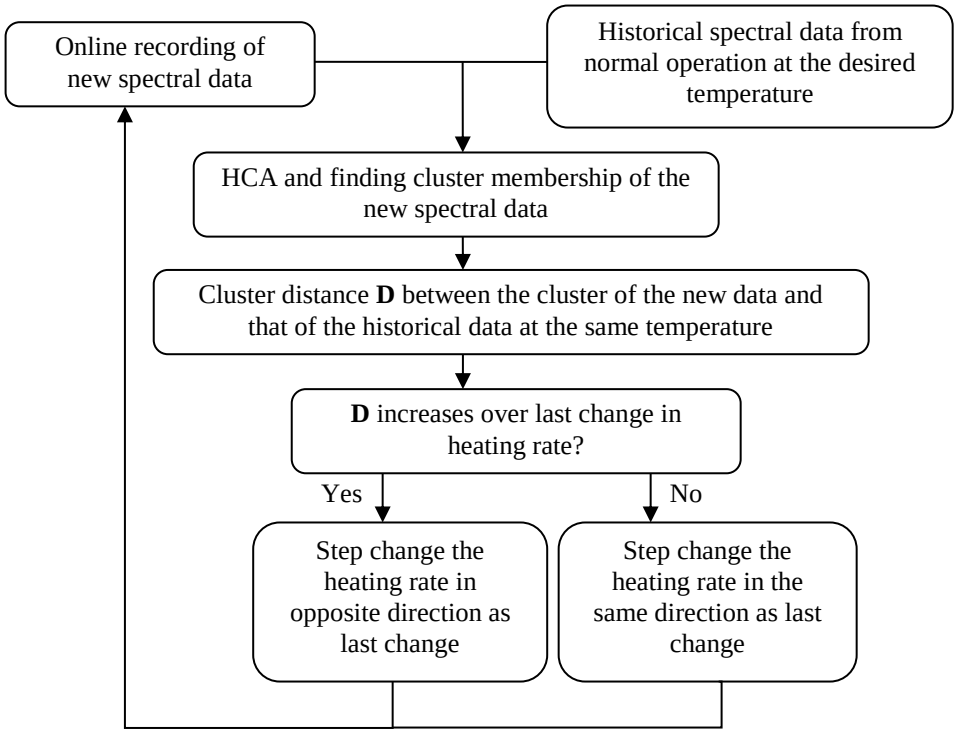


Figure 6-10. Algorithm for temperature control using HCA based on online and historical spectral data

6.5 Multivariate data analysis for analysis and system design

6.5.1 PCA and HCA for optimization of lumped kinetic modelling

As discussed in Section 6.3, PCA and HCA of the historical data could be used to design an improved lumped kinetic model. The result shown in **Figure**

6-2 suggests a lumping method based on the carbon number distribution as shown in **Figure 6-11**.

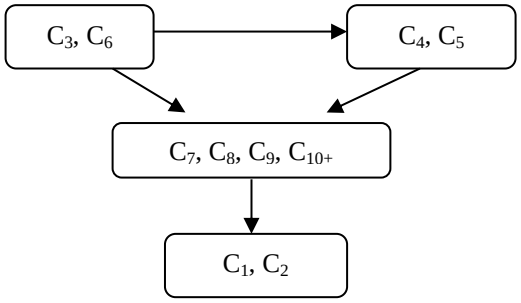


Figure 6-11. Lumping method for lumped kinetic modelling by carbon number

Future work is needed to validate this idea, and more data needs to be obtained to optimize the modelling in combination with the lumping algorithm employed in Chapter 5.

6.5.2 PCA and HCA for sample calibration

In experiments producing a large amount of samples for offline analysis, it is usually a big expense and even a waste of energy to analyze all the samples by expensive and slow methods such as offline gas chromatography, nuclear magnetic resonance and so on. HCA, combined with the fast and cheap online spectroscopy analysis, could be employed to dramatically reduce this work by clustering the samples into clusters and selecting samples from each cluster for offline analysis.

6.5.3 Multivariate data analysis for catalyst performance analysis

Catalyst deactivation is inevitable in most reaction system. This change in the catalyst performance with time will be reflected in the product species and their concentrations during over a period of time, and may result in variation in the related spectral patterns. Thus, with PCA and HCA to be used to track the change of catalyst performance with time and evaluate deactivation.

6.6 Conclusions

Multivariate data analysis including PCA and HCA is employed on the carbon number distribution data and also the online IR spectra data for analysis and design of a process control algorithm of the reaction system. The variation of the concentrations of the carbon number with the change of reaction conditions is evaluated with PCA and HCA for further understanding the catalytic conversion of propylene over ZSM-5 catalyst.

The mapping of the spectral patterns with the reaction conditions using PCA and HCA is validated and algorithms for online outlier detection and process control are proposed. Methods are also proposed to optimize the lumped kinetic modelling and to simplify the offline sample analysis. Future work needs to be done to further develop these algorithms and implement them with online spectroscopic analysis.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

To study the conversion process of pyrolysis gas from cracking of bitumen, model feed propylene was used and catalytically converted to C_1 to C_{10+} over H-ZSM-5 catalyst in a packed continuous tubing reactor. The process was tested under different feed flow rate and reaction temperature at near atmospheric pressure. A lumped kinetic model was built for the process and multivariate data analysis was employed to analyze the spectral data for process analysis and control.

As discussed in the study in the previous chapters, several conclusions are gained as below:

1) A reaction mechanism is proposed for the process of catalytic conversion of propylene over ZSM-5 catalyst which explains the distribution of carbon numbers in the reaction products well. Also, increasing the reaction temperature will affect the carbon numbers distribution leading to less conversion of propylene to heavier hydrocarbons, and meantime will increase the yield of C_1-C_2 hydrocarbons at the same WHSV. The aromatic yield will decrease with an increase in the temperature from 330 ℃ to 370 ℃ (low WHSV) or 390 ℃ (high WHSV); then it increases from 390 ℃ to 410 ℃ and again decreases up to 450 ℃. Additionally, the aromatic yield is much less sensitive to WHSV than the reaction temperature. Also, increasing the WHSV will decrease the conversion of propylene and the yield of C_{4+}

hydrocarbons, C_1-C_2 and aromatics. Thus, this process has the potential to be used for the conversion of propylene to gasoline range hydrocarbons with high material efficiency.

2) A lumped kinetic model with six lumps was built for the catalytic conversion of propylene over ZSM-5 zeolite catalyst. The kinetic parameters are estimated based on the best fit of the model. The prediction of the concentrations of the six chemical lumps agrees well with the experimental values, which indicate that the model may be used to make predictions for other experimental conditions.
3) PCA and HCA are employed to analyze the carbon number distribution data and the online IR spectral data for analysis and process control of the reaction system. The variation of the concentrations of the carbon number with the change in reaction conditions is further analyzed for understanding the catalytic conversion of propylene over ZSM-5 catalyst. Mapping of the spectral patterns with the reaction conditions is validated, and an algorithm for outlier detection and a strategy for process control are proposed on the basis of HCA of historical data. Methods are also proposed to optimize the lumped kinetic modelling and to simplify the offline sample analysis. Future work needs to be done to validate these algorithms and extend them using online spectroscopy analysis.

7.2 Future Work

As time and resources were limited in this study, potential future work is proposed below:

1) Performance of catalyst. The range of the WHSV can be further extended to explore the performance of the catalyst under various reaction conditions. Also deactivation of the catalyst can be further studied with long running time of the catalyst and for catalysis of different feed such as real pyrolysis gas from cracking of bitumen.

2) Kinetic modelling. Further work needs to optimize the lumped kinetic modelling with the help of multivariate data analysis with more analyzing data to further understanding and evaluating the process.
3) Online spectral analysis for process analysis and control. The algorithms of process analysis and control by the online analysis of the online spectra data need more practice to validate and consummate, and to be optimized for scale-up control system.

Appendix A. Setup of the GC-FID analysis for gas samples

1. Setup of the GC-FID analysis for gas samples

Table A-1. Setup of the GC-FID analysis for the gas samples

Injection volume		
Purged packed inlet	Heater temperature	200 ℃
	Total gas flow	28ml/min
	Septum purge flow	3ml/min
	Temperature procedure	See **Table A-2.**
Column	Size	10m ×200 μm ×0.5 μm
	Type	
FID	Heater temperature	250 ℃
	H_2 flow	60ml/min
	Air flow	400ml/min
	detector	Helium
	Post run	25ml/min

2. Temperature procedure of the GC at the inlet of column

Table A-2. Temperature Procedure of the GC-FID analysis for gas samples at the inlet of column

Steps	Start/ ℃	End/ ℃	Temperature increasing rate (℃/min)	Running time/min
1	70	70	0	7
2	70	250	10	18
3	250	250	0	2
4	250	70	-30	6
5	70	70	0	8

Appendix B. Setup of the GC-FID analysis for liquid samples

1. Setup of the GC-FID Analysis for liquid samples

Table B-1. Setup of the GC-FID analysis for the liquid samples

Injection volume	0.2 μL	
Purged packed inlet of the column	heater temperature	250 ℃
	Total gas flow	25ml/min
	Septum purge flow	2ml/min
	Temperature procedure	See **Table B-2.**
Column	Size	50m ×200 μm ×0.5 μm
	type	Agilent 19091S-001
FID	Heater temperature	300 ℃
	H_2 flow	30ml/min
	Air flow	400ml/min
	detector	Helium
	Makeup Nitrogen	25ml/min

Table B-2. Temperature Procedure of the GC-FID analysis for liquid samples at the injection inlet of the GC column

Steps	Start/ ℃	End/ ℃	Increasing rate (℃/min)	Running time/min
1	40	150	5	22
2	150	150	0	5

Appendix C. Selected GC-FID patterns of gas samples

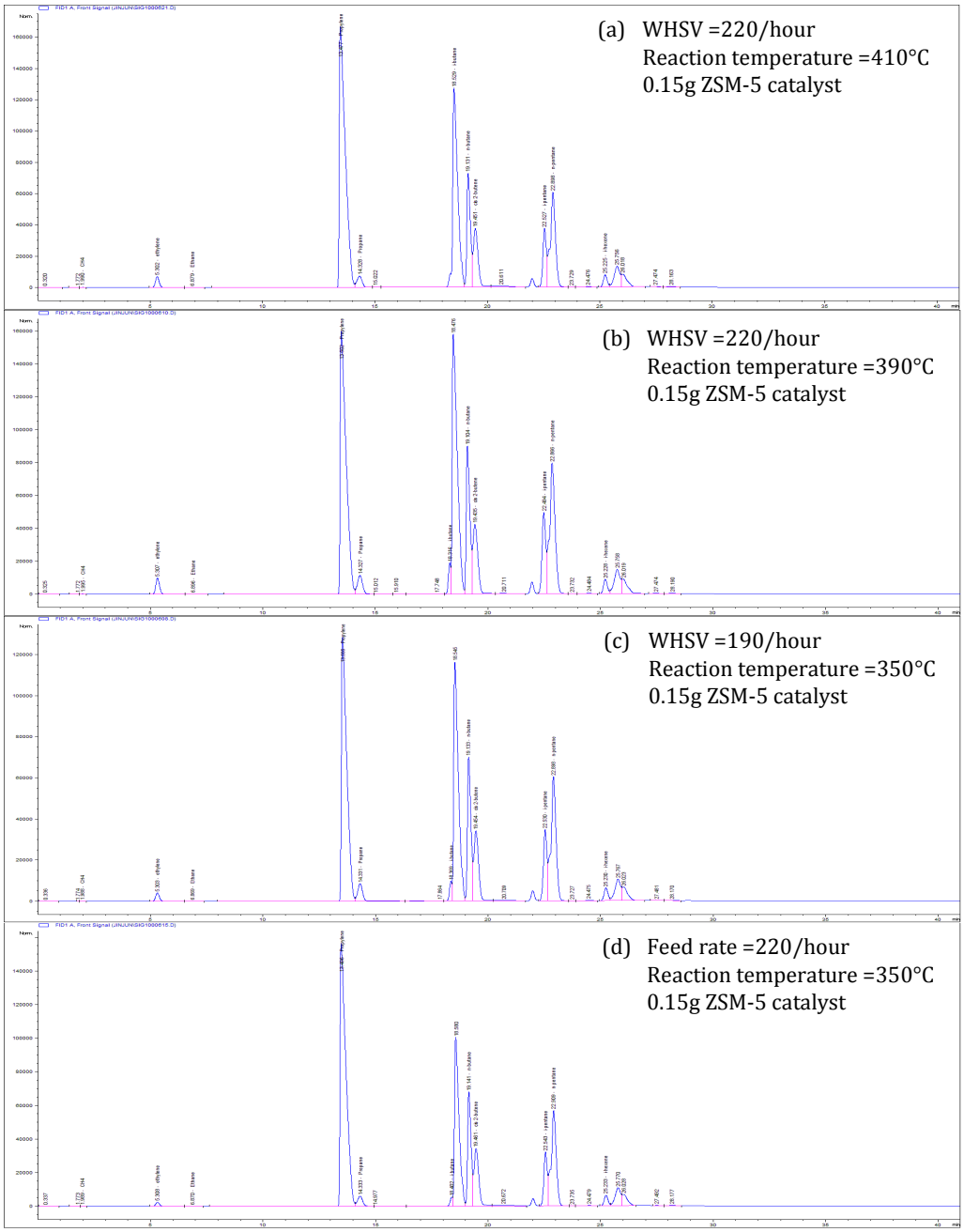


Figure C-1. GC-FID Selected patterns of gas samples from experiments under various reaction conditions

Appendix D. Selected GC-FID patterns of liquid samples

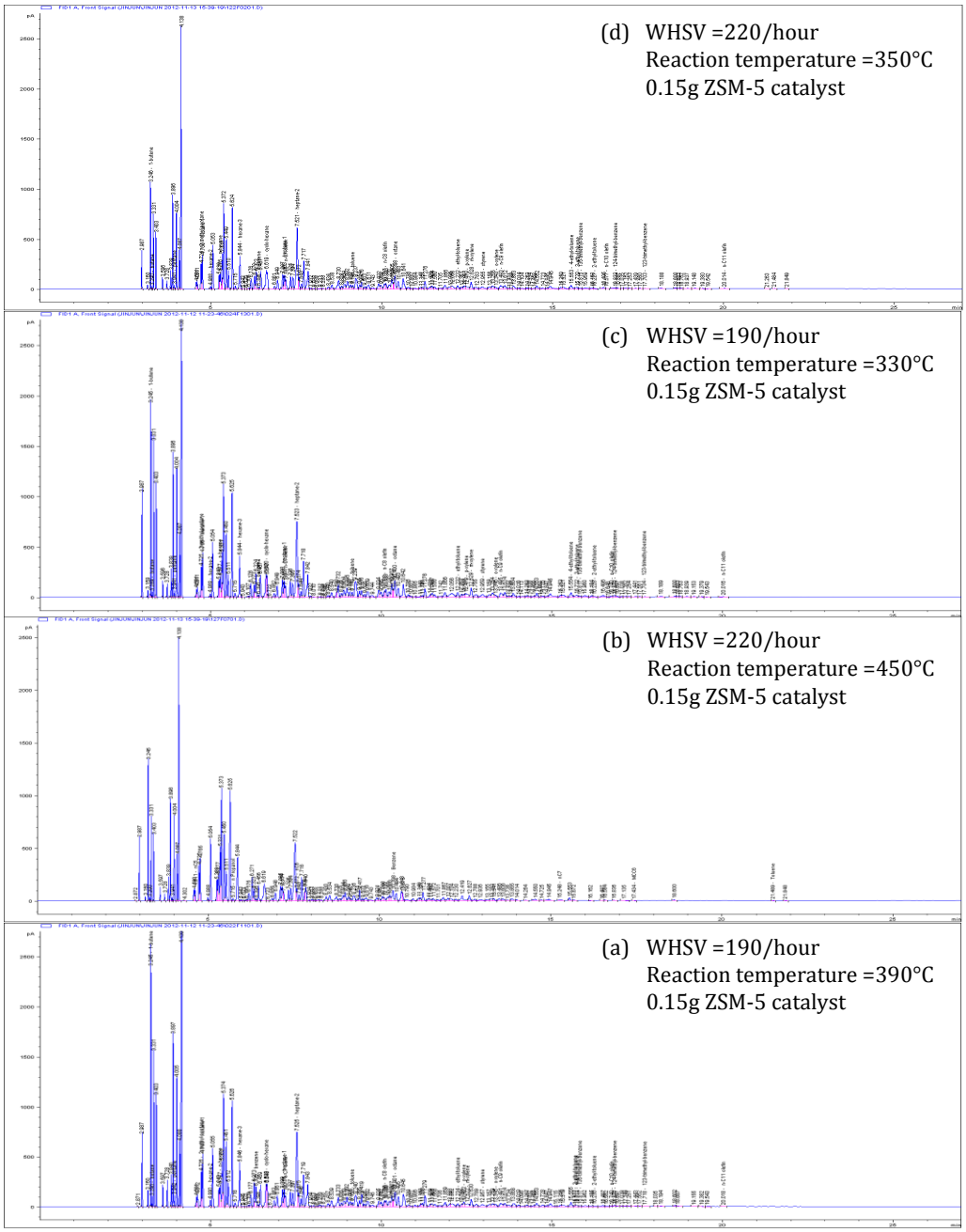


Figure D-1. GC-FID Selected patterns of liquid samples from experiments under various reaction conditions

Appendix E. MATLAB code for the lumped kinetic modelling

1. The *ccalval.m* includes a function calculating the concentrations of the lumps of chemicals in the lumped kinetic modelling under reaction conditions of various reaction temperature and feed rate.

---------- Code of *ccalval.m* begins ----------

```
function ccal = ccalval (initial )

    % cross-area of the tubing reactor, unit in square meter
    crossA =1.8374*1e-4;

   % feed volume flow, 1 atm and 20 deg, unit of m^3/sec
   % file fflow.txt contains the feed flow in unit of ml/min
    feedflow =load('fflow.txt').*1e-6/60;

    % volume of catalyst, unit of m^3
    cat_vol = 1.95*1e-6;
   % reaction temperature, unit of K
   % file temp.txt contains the reaction temperature matrix for experiments
    temp = load('temp.txt') +273.15;

   % linear flow velocity at inlet of catalysis area under the reaction temperature, 1atm
   % velocity transferred from room temperature to the reaction temperature
    feedflow = feedflow/293.15 .*temp;
    in_linearVel = feedflow/crossA;

   % inlet concentration of the feed, only determined by temperature and pressure,
   % n/V=P/RT=constant
    c_inlet = 1000/22.41 .*(293.15 ./temp);
    R = 8.3145;                    % gas constant, unit of J/(K.mol)
    [mt nt] = size(temp);          % size of temperature matrix

  % the length of the catalyst in the reactor along the flow direction
    lmax = cat_vol/crossA;

  % rate constant by activation energy, temperature and rate constant coefficient
    rea = initial;             % initial value for the reactions kinetic parameters
    ccal = zeros(mt,6);  % initialize the value of the calculated kinetic parameters
    global k0;                 % reaction rate constant
    global vbeta;
  % divide the catalysis length into 500 units for differential calculation
       ii =500;
       dx =lmax/ii;

    % for each experiment, concentration of chemical lumps is caculated
   for i =1:mt;
     % initial concentration of the chemical lumps, con21 for the propylene feed
```

```
        con11 =0;
        con21 =c_inlet(i);
        con31 =0;
        con41 =0;
        con51 =0;
        con61 =0;
        vbeta11 =in_linearVel(i);
    % Arrhenius equations for the reactions, all the parameters rescaled
        k0(1) = exp(rea(1)-20000*rea(2)/R/temp(i));
        k0(2) = exp(rea(4)-20000*rea(5)/R/temp(i));
        k0(3) = exp(rea(7)-20000*rea(8)/R/temp(i));
        k0(4) = exp(rea(10)-20000*rea(11)/R/temp(i));
        k0(5) = exp(rea(13)-20000*rea(14)/R/temp(i));
        k0(6) = exp(rea(16)-20000*rea(17)/R/temp(i));
        k0(7) = exp(rea(19)-20000*rea(20)/R/temp(i));

    % ODE equations by differential equation arrays for the reaction rate expression
        for ij=1:ii;
            % sum of reaction rates for the whole reaction network
            rsum =-0.488*k0(2)*con21^rea(6) -0.305*k0(4)*con41^rea(12)+ …
                    +1.036*k0(3)*con41^rea(9) +0.952*k0(6)*con51^rea(18) + ...
                      +0.864*k0(7)*con51^rea(21)  +1.929*k0(5)*con51^rea(15);
            crsum =rsum*0.024224*temp(i)/293.15;
            vbeta21 =vbeta11 +rsum*0.024224*temp(i)/293.15*dx;
            con12 =con11 +1/vbeta11*(2.036*k0(3)*con41^rea(9) +
                    2.99*k0(4)*con51^rea(12))*dx;
            con22 =con21 +1/vbeta11*(-k0(1)*con21^rea(3) -k0(2)*con21^rea(6))*dx;
            con32 =con31 +1/vbeta11*(k0(1)*con21^rea(3) +
                    1.864*k0(5)*con51^rea(15))*dx;
            con42 =con41 +1/vbeta11*(-k0(3)*con41^rea(9)+
                    1.439*k0(6)*con51^rea(18))*dx;
            con52 =con51 +1/vbeta11*(0.512*k0(2)*con21^rea(6) - k0(4)*con51^rea(12) -
                    k0(5)*con51^rea(15) - k0(6)*con51^rea(18) -k0(7)*con51^rea(21))*dx;
            con62 =con61 +1/vbeta11*(0.781*k0(7)*con51^rea(21))*dx;
         % continuity of the outlet and inlet of successive differential catalyst units
            con11 =con12;
            con21 =con22;
            con31 =con32;
            con41 =con42;
            con51 =con52;
            con61 =con62;
          % residence time integration by differential
            tt(i) =tt(i) +dx/vbeta11;
            vbeta11 =vbeta21;
        end

    % get the last concentrations of the lumps in line as the outlet concentrations
        clastone =[con12 con22 con32 con42 con52 con62];
        ccal(i,:) =clastone;
    end
end
```

----------End of *ccalval.m*----------

2. The file *minobj.m* containing a function first calls and transfers the initial guess of the kinetic parameters to *ccalval.m* to calculate the concentration of the chemical lumps at the reactor outlet by the lumped model, and then load the experimental data of the concentration to calculate the difference of them as the output of itself.

----------Code of *minobj.m* begins----------

```
function ffmin = minobj( initialrea )
   % c_utlet.txt contains the experimental data of concentration of chemical lumps
    c_outlet =load('c_outlet.txt');

   % calculate the concentration of the lumps by the lumped model
   ccal =ccalval(initialrea);

   % min objective function
    fmin =sum(((c_outlet-ccal)./c_outlet).^2);
    ffmin =sum(fmin);
end
```

----------Code of *minobj.m* ends----------

3. The file *minimize.m* takes the output from *minobj.m* as objective and calls *fmincon* to guess the value of kinetic parameters to minimize the objective value.

----------Code of *minimize.m* begins----------

```
function [ rea_est,ffval ] = minimize(initial)

   % constrains on the kinetic parameters
    Aeq =[];
    beq =[];
    lb =zeros(21,1);
    ub =zeros(21,1);
    for i=1:7;
       lb(3*i-2) =0.5;
       lb(3*i-1) =0.3;
       lb(3*i) =1;
       ub(3*i-1) =20;
       ub(3*i-2) =10;
       ub(3*i) =2;
    end
```

```
    options =optimset('Algorithm','interior-
                point','Display','iter','TolFun',1e30,'TolCon',1e-30,'TolX',1e-
                50,'MaxFunEvals',24000);
    [rea_est,ffval] =fmincon(@minobj,initial,[],[],Aeq,beq,lb,ub,[],options);

end
```

----------Code of *minimize.m* ends----------

Appendix F. Related MATLAB code for PCA and HCA

Range selection from the original spectra data
The function *rangecut.m* selects the desired range of wavenumber data from the original spectra data.

----------Code of *rangecut.m* begins----------

```
function [newwavenum, newcutting] = rangecut( xlsmatrix,n,soilt1,end1)

 [wavenum irband] =xlsdeal(xlsmatrix);
 cutting =zeros(n,2);
 [m1,m2] =size(wavenum);
 for i=1:n;
    for j=1:m1-1;
      if wavenum(j+1)>=soilt1(i)&&wavenum(j)<=soilt1(i)
         cutting(i,1)=j;
      end
      if wavenum(j+1)>=end1(i)&&wavenum(j)<=end1(i)
         cutting(i,2)=j;
      end;
    end;
 end;

 for ii=1:n;
    if ii==1;
       st1(ii) =1;
    else
       st1(ii) =sum(cutting(1:ii-1,2))-sum(cutting(1:ii-1,1))+ ii;
    end;
       en1(ii) =sum(cutting(1:ii,2))-sum(cutting(1:ii,1))+ ii;
 end;

 for ii =1:n;
    newwavenum(st1(ii):en1(ii)) =wavenum(cutting(ii,1):cutting(ii,2));
    newcutting(st1(ii):en1(ii),:) =irband(cutting(ii,1):cutting(ii,2),:);
 end;
end
```

----------Code of *rangecut.m* ends----------

Data normalization
The function *normalize.m* scale the data to the range of [0,1].

----------Code of *normalize.m* begins----------

```
function nordata = normalize( matdata )

  [m n]=size(matdata);
  nordata =matdata;
```

```
  mincol =min(matdata);
  maxcol =max(matdata);

  for i=1:n;
     for j=1:m;
        nordata(j,i) =(matdata(j,i)-mincol(i))/(maxcol(i)-mincol(i));
     end
  end

end
```

----------Code of *rangecut.m* ends----------

Main code for PCA

```
----------Code of PCAprocess.m begins----------
function [ prinnum cutvariance cutscore  ] = PCAprocess( matdata, cutvalue)
   matdata =normalize(matdata);
   [coefs,scores,variances,t2] = princomp(matdata)
   percent_variance =variances/sum(variances);
   [mv nv] =size(percent_variance);

   if cutvalue<=percent_variance(1)
         prinnum =1;
   end
   if cutvalue>percent_variance(1)
      for i =1:nv-1;
if cutvalue>sum(percent_variance(1:i)) && cutvalue<=sum(percent_variance(1:i+1))
               prinnum =i;
          end
      end
   end
end
```

----------Code of *PCAprocess.m* ends----------

Main code for HCA

```
----------Code of HCAprocess.m begins----------
function [h t] = HCAprocess( matdata )

  matdata =normalize(matdata);
  distance =pdist(matdata');
  zlink =linkage(distance, 'ward');
  [h t] =dendrogram(zlink);

end
```

----------Code of *HCAprocess.m* ends----------