ANALYSIS OF PUBLISHED DATA IN THE LITERATURE TO EXTRACT KNOWLEDGE FOR FISCHER-TROPSCH SYNTHESIS

by

Barış Burnak

B.S., Chemical Engineering, Boğaziçi University, 2013

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering

Boğaziçi University

2015

to my muse,

Metin Burnak

ACKNOWLEDGEMENTS

Before presenting my study, I would like to take this brief moment to express my gratitude to those who supported and inspired me during this enervating period.

Foremost, I would like to thank my thesis supervisor, Prof. Ramazan Yıldırım, for his farsighted guidance and encouraging enthusiasm to produce fruitful results even from the slightest piece of information. His dedication and patience with my thesis always galvanized me to perform my best. I feel very lucky for having the chance to work with him.

I also thank my thesis committee, Assoc. Prof. Ahmet Kerim Avcı and Assist. Prof. Mehmet Erdem Günay, for sparing their valuable time contributing to my study.

I profoundly thank Elif Can for accompanying me in this long challenging journey. She kept me away from being lost in the details throughout my studies; and I would most probably be struggling with constructing new models as the reader peruses these lines.

Unequivocal contribution of Elif Erdinç and Serhat Erşahin, especially in my early studies constituted the fundamental basis of my knowledge on terminology of catalytic experiments. I owe them for sharing any kind of information with joy.

I would like to thank my dearest friends Begüm Alaybeyoğlu, Özgür Yaşar Çağlar, Salih Emre Demirel, Coşar Doğa Demirhan, Elif Gençtürk and Manouchehr Nadjafi for being awesome people in my life. I have never faced the brutal ruthlessness of the process of writing a thesis with their support. I also want to thank Bilge Kerem Aksakal for helping me unconditionally whenever I need, and not withholding his wisdom from me.

I would like thank all the members of the CATREL family, and also Yakup Bal and Murat Düzgünoğlu for maintaining a perfect working environment.

I specially thank to my beloved sister and mother, Oya and Kadriye Burnak, for their support throughout my entire life. Their trust has encouraged me to pursue my dreams.

ABSTRACT

ANALYSIS OF PUBLISHED DATA IN THE LITERATURE TO EXTRACT KNOWLEDGE FOR FISCHER-TROPSCH SYNTHESIS

In this study, empirical global models for Fischer-Tropsch Synthesis (FTS) performance were constructed by using machine learning algorithms on the experimental data published in the literature. CO conversion in FTS was modelled as a function of catalyst design variables, physical properties of the catalyst, and operating conditions by using multiple linear regression, artificial neural network (ANN), decision tree, and random forest in R 3.1.2 environment. The missing values in the physical properties were estimated by constructing ANN models on the known instances. Multiple linear regression and decision tree classification models on the completed dataset yielded the least reliable results with low prediction accuracies. The predictive power of the ANN and random forest models, on the other hand, were much higher with R^2 values of 0.40 and 0.59, respectively. The importance analysis on both models revealed that operating temperature is the most crucial attribute in FTS catalytic activity. The predictive power of the models were improved by constructing sub-models trained by the Co-based, Fe-based, low temperature, and high temperature subsets, which were generated by splitting the entire dataset with respect to the base metal and operating temperatures. The R^2 values for these sub-models ranged between 0.43 and 0.53 with ANN, and 0.62 and 0.66 with random forest algorithms. According to the importance analyses on these sub-models, the catalyst design variables were determined to be more dominant in Fe-based subset, whereas operating conditions had the utmost importance for Co-based subset. Sensitivity analysis was conducted to observe the effect of three of the most important attributes in the ANN models. The residual analysis on each model revealed that all of the algorithms failed to represent the effect of time on stream on CO conversion. In addition to these predictive approaches, principal component analysis was used to describe the correlations between the input variables and CO conversion. It validated the outcomes of ANN and random forest models by determining the operating temperature as the most important attribute.

ÖZET

FISCHER-TROPSCH SENTEZİ İLE İLGİLİ BİLGİ ÇIKARIMI İÇİN LİTERATÜRDE YAYIMLANAN VERİLERİN ANALİZİ

Bu çalışmada, literatürde yayımlanan deneysel veriler üzerine özdevimli öğrenme algoritmaları kullanılarak Fischer-Tropshc Sentezi (FTS) üzerine genelgeçer ampirik modeller oluşturulmuştur. FTS'deki CO dönüşümü R 3.1.2 ortamında çoklu doğrusal regresyon, yapay sinir ağı (YSA), karar ağacı ve rastgele orman (random forest) algoritmaları kullanılarak katalizör tasarımı değişkenleri, katalizörün fiziksel özellikleri ve operasyon koşullarının bir fonksiyonu olarak modellenmiştir. Katalizörün fiziksel özelliklerindeki eksik değerler, bilinen değerlere dayanan YSA modelleri kullanılarak tahmin edilmiştir. Tamamlanan veri tabanı üzerine kurulan çoklu doğrusal regresyon ve karar ağacı sınıflandırma modelleri deneme seti üzerinde düşük tahmin oranıyla en az güvenilir sonuçları üretmişlerdir. Öte yandan, YSA ve rastgele orman modellerinin tahmin gücü sırasıyla 0.40 ve 0.59 R^2 değerleri ile çok daha yüksek bulundu. Bu iki model üzerinde yapılan önem analizi, operasyon sıcaklığının FTS katalitik aktivitesindeki en önemli değişken olduğunu göstermiştir. Bu modellerin tahmin güçleri, bütün veri tabanının temel metal ve operasyon sıcaklığına göre ayrılmış Co tabanlı, Fe tabanlı, düşük sıcaklık ve yüksek sıcaklık alt gruplarıyla eğitilmiş alt modeller oluşturularak geliştirilmiştir. Bu alt modellerin R^2 değerleri YSA algoritmasıyla 0.43 ve 0.53, rastgele orman algoritmasıyla ise 0.62 ve 0.66 değerleri arasında değişkenlik göstermiştir. Bu alt modellerde yapılan önem analizi, Fe tabanlı alt grupta katalizör hazırlama değişkenleri daha baskınken, Co tabanlı alt grupta operasyon koşullarının en önemli değişkenler olduklarını göstermiştir. YSA modellerindeki en önemli üç değişkenin etkisini gözlemlemek için hassaslık analizi yapılmıştır. Bütün modeller üzerinde yapılan kalıntı analizi, kullanılan algoritmaların akıştaki zamanın CO dönüşümü üzerindeki etkisini yansıtamadığını göstermiştir. Bu kestirimci yaklaşımlara ilaveten, giriş değişkenleri ve CO dönüşümü arasındaki korelasyonları betimlemek için temel bileşen analizi kullanılmıştır. Bu analiz, operasyon sıcaklığının en önemli değişken olduğunu belirleyerek YSA ve rastgele orman modellerinin sonuçlarını doğrulamıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

f	Activation function
R^2	Coefficient of determination
w	Weight of the connection
x_j	j^{th} input variable of an instance
y	Output of an instance
β	Coefficient of multiple linear regression
ε	Residual between actual and predicted output values
θ	Bias

LIST OF ACRONYMS/ABBREVIATIONS

AC	Activated carbon
ANN	Artificial neural network
CART	Classification and regression tree
CB	Carbon black
CMC	Carbon micro-coils
CNF	Carbon nanofiber
CNT	Carbon nanotube
FTS	Fischer-Tropsch Synthesis
GHSV	Gas hourly space velocity
HDP	Homogeneous deposition precipitation
HP	Homogeneous precipitation
HTFT	High temperature Fischer-Tropsch Synthesis
IWI	Incipient-to-wetness
LTFT	Low temperature Fischer-Tropsch Synthesis
MAE	Mean absolute error
ME	Microemulsion
OMC	Ordered mesoporous carbon
PCA	Principal component analysis
pHP	pH precipitation
PR	Precipitation
PVI	Pore volume impregnation
RMSE	Root mean square error
SD	Spray-drying
SI	Slurry impregnation
TOS	Time on stream
WGS	Water gas shift reaction
WI	Wetness impregnation

1. INTRODUCTION

The rapidly increasing human population causes an increase in energy consumption in the world. Although the market share of crude oil in energy supply has been decreasing for the last fifteen year, it is still the leading source of energy as of 2015 [1]. Depleting crude oil reserves is terminating the era of cheap liquid fuels [2], steering the government policies, companies, and academics towards alternatives sources of energy, and more efficient utilization of the existing sources.

Fischer-Tropsch Synthesis (FTS) is an alternative route to produce liquid fuels from coal, natural gas, shale gas, coal-bed gas and biogas feedstocks, and it is known for almost a century. Since direct transformation of these energy sources to liquid fuels is still expensive and impractical, FTS has been attracting renewed interest in the last decade. FTS products include a wide spectrum, including long hydrocarbon chains such as gasoline, diesel, jet fuel, and also precursors of fine chemicals, such as lower olefins. In addition to the hydrocarbons, FTS can yield organic oxygenates such as methanol, ethanol and higher alcohols. Producing these products from the aforementioned sources is a two-step process; gasification or reforming of the raw material, and hydrogenation of the produced synthesis gas. Employing a two-step production technique minimizes the sulfur and nitrogen content in the products, which is a valuable outcome because of the stringent environmental regulations enforced by the governments [3]. Additionally, FTS provides independence from crude oil imports and fluctuating oil prices [4]. New incentives on FTS technology have been introduced in the last two decades both in developed and developing countries such as USA, China, Qatar, Malaysia, Japan, and Italy [5].

Designing and developing novel catalyst with high activity and preferred selectivity; and operating under the optimum conditions is the main focus of the academic community in the field. Although activity and stability of the catalyst are utterly crucial in FTS process, the academics mainly lean on tuning the product selectivity. A typical FTS process produces hydrocarbon chains of various lengths, methane, and carbon dioxide. Methane is the most undesired product,since it is in gaseous phase in normal conditions, and thus the least practical mean for energy storage and transportation. Accordingly, the studies focus on

maximizing the selectivity for hydrocarbons chains longer than five carbons (C_5^+) to reduce the operating costs for downstream separation [3].

Typical active metals used in FTS catalysts are Co, Fe, Ru, and Ni. Fe and Co are heavily used in industrial applications because they are inexpensive, and widely available in the market. However, various metal promoters are required to improve their activity, selectivity, and stability. Fe-based catalysts are mostly used in coal-to-liquid and biomass-to-liquid processes, which operate with hydrogen poor feedstocks due to the nature of the raw material. They can operate under wide ranges of operating temperatures and syngas compositions without significantly changing the catalytic activity and selectivity of the products. Co-based catalysts, on the other hand are more expensive in the market, yet they are more prone to produce longer hydrocarbon chains, such as diesel fuel and wax. Ru is the most active and efficient metal for FTS process among other types of catalysts. It is highly active and selective to long hydrocarbon chains even at low operating temperatures without any promoting metals. However, its scarcity and high price in the market limits its widespread utilization in the industrial applications. Consequently, the studies in the literature mostly focus on improving the performance of Co and Fe based catalysts by adding on various supports, promoting with different types of metals, modifying the textural properties by trying numerous metal addition techniques, and tuning the calcination and reduction conditions [6].

Considering that FTS is invented in 1920s, there is a vast amount of improvement and research accumulated in the literature over years. Abundance of the attributes related to the reaction makes the system involved and recondite. Hence, data driven models can be constructed to analyze complete past studies on the topic, in addition to dealing with this complexity by analyzing physical and chemical relations.

Machine learning algorithms, at this point, can alleviate the burden of the researchers in the field by narrowing the scope down to the potentially fruitful areas. These algorithms use example data or past experience to construct a parsimonious model to predict future events [7]. Numerous machine learning algorithms were developed for data mining purposes throughout the years. Gunay and Yildirim implemented these algorithms successfully to predict the outcomes of catalytic performance [8] and reaction kinetic studies [9].

In this study, CO conversion in FTS was modelled empirically by using the published articles in the literature and implementing machine learning algorithms to extract various kinds of useful knowledge. In Chapter 2, a thorough literature survey on FTS and machine learning algorithms used in this study were presented. The details of the database constructed by the articles published in the last decade, and the implementation of the algorithms were described in Chapter 3. The obtained result and detailed discussion on the outcomes were presented in Chapter 4. Finally, the conclusions of this study and recommendations for future work were provided in Chapter 5.

2. THESIS BACKGROUND

2.1. Fischer-Tropsch Synthesis

Fischer-Tropsch Synthesis (FTS) is a clean and efficient method to produce sulfur and nitrogen-free fuels and various oxygenates from syngas derived from natural gas, coal, or biomass [10]. In FTS, carbon monoxide is hydrogenated by the following catalytic surface polymerization reactions [11]:

$$2n\ H_2 + nCO \rightarrow\ C_nH_{2n} + nH_2O \tag{2.1}$$

$$(2n+1)H_2 + nCO \rightarrow C_nH_{2n+2} + nH_2O \tag{2.2}$$

$$2n\ H_2 + nCO \rightarrow C_nH_{2n+1}OH + (n\text{-}1)H_2O \tag{2.3}$$

Reactions given above are highly exothermic, and their enthalpies are determined by the chain length of the end product. As an exception to the stated desired reactions, Equation 2.2 can yield methane formation reaction, which is one of the main challenges that limit the FTS efficiency. Most of the FTS performance studies basically focus on maximizing selectivity to higher hydrocarbon chains, while minimizing the rate of methane formation. Other undesired reactions are mainly Boudouard reaction, formation of bulk carbides and carbonaceous materials [12]. Water-gas-shift (WGS) is another undesired reaction for hydrogen rich syngas feedstocks, whereas hydrogen poor syngas feedstocks require WGS parallel with FTS to increase hydrogen to carbon monoxide ratio [13].

Conventionally, in industrial practices, the reaction temperature is maintained between 330 – 350 °C for high temperature Fischer-Tropsch (HTFT) conditions to increase gasoline and light olefin yield; low temperature Fischer-Tropsch (LTFT) is performed between 220 – 250 °C to increase wax and diesel fuel production. A typical FTS reactor operates under 25 – 45 bars [6].

2.1.1. Effect of Active Metal

FTS is typically carried out on metal oxide supported Co and Ru catalysts, as well as unsupported Fe based catalysts. Yet, on-going research is focused on understanding the function of different chemical states of these active phases for improved catalytic performances [3].

Cobalt is designated as the most promising active metal for CO hydrogenation for hydrogen rich feedstocks due to its high chain growth probability (α), low deactivation rate, low WGS and relatively low price [14, 15].

Cobalt loading is one of the most important parameters affecting FTS activity of catalyst activity. Increase in the amount of cobalt loading leads to high CO conversion, low selectivity to CH_4, and high selectivity to long hydrocarbon chains due to agglomeration of cobalt crystallites [16]. Trépanier *et al.* conducted FTS at 220 °C, 2.0 MPa in a fixed-bed microreactor by using multi-walled carbon nanotube (MWCNT) supported cobalt catalyst to investigate the effect of cobalt loading. CO conversion increased as the cobalt loading is incremented from 15 wt.% to 30 wt.%. Doubling the cobalt loading increased the number of available active sites and hence, improved FTS rate by 70 %. [15]. Khodaei *et al.* observed a similar increasing trend in CO conversion on a (Co-Mn)/TiO_2 catalyst. They reported that conversion increases at 260 °C and 0.1 MPa, as cobalt – manganese weight fraction increased from 5 wt.% to 30 wt.%. Yet, further increase in the cobalt loading decreases the specific surface area, pore diameter and pore volume of the catalyst, which yields lower conversion levels [17].

However, since the main goal of FTS is mostly to maximize C_5^+ yield, it is required to tune hydrocarbon selectivity values along with CO conversion level. Studies pointed out that increasing Co loading suppresses methane formation and increases C_5^+ selectivity [16-18].

Selecting the correct Co precursor has the utmost importance to obtain a high performance FTS catalyst. Yu *et al*. and Bae *et al*. focused on nitrate, acetate, and chloride precursors and reported conflicting results on the superior precursor. The former claimed higher CO conversion and C_5^+ selectivity is attained with Co acetate, whereas the latter

suggested Co nitrate gives better catalytic activity. However, both of the research groups agreed that higher reducibility and dispersion levels lead to higher C_5^+ yield [19, 20].

Iron based FTS catalysts are commercially used with hydrogen-deficient syngas feedstocks, since they greatly enhance WGS reaction [21].

Functions of the active sites and phases of the iron based catalysts are still unclear and yet to be discovered. Nevertheless, recent studies indicated that FTS activity mostly occurs on iron carbides, instead of metallic iron, whereas iron oxides are responsible for the high WGS reaction rates [22].

On-going research both focus on supported and unsupported Fe-based catalysts. Galvis *et al*. compared performances of unsupported base iron with α-Al_2O_3 supported iron catalyst. They reported that supported iron yields better stability along with higher CO conversion and C_5^+ selectivity [23]. A similar study was conducted by Baranak *et al*., comparing unsupported base iron with ZSM-5 supported iron catalyst and physical mixture of the two catalysts. They suggested that productivity of the base iron is much more limited due to its low surface area and pore volume. Additionally, investigation of the iron loading revealed that increasing iron load from 9 wt.% to 18 wt.% increases CO conversion, and decreases CO_2 and CH_4 formation [24].

The type of the precursor predetermines carbide formation and hence, FTS activity of the Fe-based catalysts. Kang *et al*. compared acetate, nitrate, and citrate iron precursors on Al_2O_3 and SiO_2 supports at 300 °C and 1.0 MPa. Acetate precursor was found to be the least active precursor irrespective of the support type. In addition; it yielded significantly higher CH_4 selectivity on both support types due to unreduced iron oxides [21].

Ruthenium based catalysts are known for their high FTS activity, but high cost and limited availability of ruthenium precursors in the market inhibits their commercial use. Ruthenium chloride is a relatively cheap precursor and may be a suitable candidate for the ruthenium based FTS catalysts. Nurunnabi *et al*. studied on ruthenium based catalysts prepared by its chloride precursor. CO conversion increased with increasing Ru loading between 1 – 5 wt.% at 220 °C and 20 bar, but decreased at 10 wt.% Ru. In addition to high

CO conversion levels, low CH_4 selectivity and high C_5^+ selectivity for all Ru loadings pointed out the superior performance of ruthenium based catalysts [25].

Co-Fe bimetallic catalysts have been a subject numerous studies to combine high dispersion, reducibility, and stability qualities of cobalt with economic advantage of iron. Griboval-Constant *et al*. tried to optimize the performance of such a bimetallic FTS catalyst, and compared those results with monometallic supported cobalt and iron catalysts. Although 10wt.%Co/Al_2O_3 gave higher CO conversion, higher C_5^+ selectivity, and lower CO_2 selectivity; 5wt.%Co-5wt.%Fe/Al_2O_3 yielded higher selectivity to olefins at 240 °C and 20 bar. The decrease in C_5^+ selectivity in bimetallic catalysts were explained by formation of mixed Co-Fe compounds on the support surface. Additionally, high dispersion of Co in the presence of Fe species resulted in high CH_4 selectivity [26].

2.1.2. Effect of Support

An effective FTS catalyst design requires sufficient resistance to attrition and efficient heat dissipation. Additionally, chemical interaction between the active phase and the support should be mild. If the interaction is too weak, the active phase will fail to disperse on the surface. On the other hand, too strong interaction will decrease the extent of reduction of the active phase. Moreover, physical properties of the support such as surface area, pore volume and pore diameter significantly affects FTS activity since they influence the size and specific surface area of the active phase, diffusion of the reactants and products [6].

Most of the Co and Ru based FTS catalysts investigated in the literature are supported with numerous metal oxides such as SiO_2, Al_2O_3, TiO_2. However, the literature has presented contradicting results on CO hydrogenation activity so far. Song *et al*. and Ngamcharussrivichai *et al*. maintained that Co is more active on SiO_2 than on Al_2O_3, both in terms of CO conversion and C_5^+ productivity [27, 28]. On the contrary, Yu *et al*. suggests that Co/Al_2O_3 yields higher CO conversion compared to Co/SiO_2 [19].

Nurunnabi *et al*. conducted a similar study with Ru as the active metal, and compared FTS activity on Al_2O_3 and SiO_2. They revealed that large particle size and high reducibility of Ru on Al_2O_3 increases CO conversion and C_{21}^+ selectivity, compared to SiO_2 support [25, 29].

The ambiguity of the nature of the active phase in Fe-based catalysts diverges the research into supported and unsupported catalysts design. Arsalanfar *et al*. focused on comparison of FTS productivity on supported and unsupported Fe-based catalysts. CO conversion was lower on unsupported catalyst compared to TiO_2, SiO_2, and MgO supports, but higher than Al_2O_3 support. SiO_2 yielded the highest CO conversion. In addition, it was found that unsupported catalyst produced high amounts of methane compared to supported catalysts [29]. Nevertheless, Kang *et al*. presented another contradicting result indicating that CO conversion on Al_2O_3 is much higher than on SiO_2. They also found that methane and CO_2 formation rates were higher and C_5^+ selectivity is lower on Al_2O_3 [21].

Other types of metal oxides tested as FTS supports in the literature are different phases of Al_2O_3 [30, 31], MnO_2 [32], ZnO [33-35], ZrO_2 [36-39], CeO_2 [40, 41], and La_2O_3 [42, 43].

The major challenge of FTS process is to produce selective production of hydrocarbons with a narrow carbon number distribution. Zeolite involved bifunctional FTS catalysts are good candidates to suppress formation rate of methane and infinitely long hydrocarbon chains due to their cracking properties. Zeolites can be either physically admixed with conventional FTS catalyst in one reactor by different configurations, coated on Co agglomerates, or act as a support on Co particles [44].

Sartipi *et al*. compared three zeolite categories mentioned above with HZSM-5 and Co/SiO_2. They physically mixed HZSM-5 with Co/SiO2, coated HZSM-5 on Co particles, and impregnated Co on HZSM-5 supports to investigate their performances. They reported a detailed study on the effect of zeolite on CO conversion and product selectivity at different FTS operating conditions. Coating HZSM-5 on Co agglomerates shifted the product selectivity towards C_2-C_4 interval from C_5^+ while decreasing CO conversion at all operating conditions. Compared to the reference Co/SiO_2 catalyst, HZSM-5 supported Co significantly increased CO conversion and shifted product selectivity from C_{12}^+ to C_2-C_{11} interval as desired. An unanticipated advantage of these two catalyst designs was decrease in CO_2 selectivity. However, physically admixed Co/SiO_2 + HZSM-5 catalyst yielded lower CO conversion and roughly same product selectivity with reference catalyst [44].

Baranak *et al*. conducted a similar study with Fe catalysts on ZSM-5 zeolite. Both physically admixed catalyst and zeolite supported iron catalyst yielded higher CO conversion with a narrower hydrocarbon chain length around C_5-C_{11} interval at 280 °C and 19 bar. Similar to Sartipi *et al*., they lowered selectivity to CO_2 with zeolite supported and physically admixed catalysts [24].

Recently developed carbon based supports have been a subject of FTS catalyst design due to their numerous advantages. Carbon materials are widely available in the market at reasonable prices, they can be regenerated by a simple oxidation process, and they can be easily tailored to adjust surface acidity [45].

Carbon nanofibers (CNF) and carbon nanotubes (CNT) have been the most widely investigated carbon based FTS support within the last decade. Liu *et al*. synthesized CNT-decorated α-Al_2O_3 supported Co catalyst to compare its physical properties and performance with Co/α-Al_2O_3 catalyst. They reported significantly higher surface area on the CNT-decorated support, which resulted in higher CO conversion. Additionally, larger metallic Co particle size on the surface increased selectivity in C_2-C_4 interval [46, 47].

Electronic structures of the carbon based supports can be easily modified to prepare a high performance FTS catalyst. Fu *et al*. doped nitrogen functional groups within CNTs. They reported that Co on N-doped CNTs yield higher CO conversion compared to base CNTs [48]. Furthermore, physical structures of CNT, such as surface area, pore volume and pore size, were modified through changing the time for acid treatment. Increasing the time for acid treatment also affected Co particle size and dispersion on the support, which affect CO conversion and product selectivity [49].

Xiong *et al*. extensively studied Co activity on carbon based support structures, especially CNT, carbon nanofibers (CNF), and carbon micro-coils (CMC). They proposed that when the calcined catalyst is reduced at 300 °C, CNT supported catalyst yields slightly higher CO conversion and significantly higher C_5^+ selectivity than on CNF and CMC. However, when the calcination temperature was 400 °C, CMC supported catalyst yielded a very high chain growth probability value with very low CO conversion. On the other hand,

high reduction temperature almost doubled CO conversion at the same reaction conditions while increasing methane formation significantly [50].

While earlier FTS studies had extensively investigated activated carbon (AC), carbon black (CB), and glassy carbon [45], recent studies have focused on ordered mesoporous carbon (OMC) [51], graphite, oxidized diamond [52], carbon spheres [53].

2.1.3. Effect of Promoters

Cobalt and iron based FTS catalysts usually require promoters to adjust the size and dispersion of the active phase. Alkali metals ions, noble metals, and transition metals are used to obtain improved FTS activity. Ruthenium based metals, on the other hand, can work properly for C_5^+ productivity without any promoters [6].

Most of the Fe-based catalysts in the literature are promoted with alkali metals, especially with Li, Na, and K. These metals function as electronic promoters, enhancing the chemisorption of CO on the active sites and inhibiting that of H_2 [6]. Although the exact nature of the promoting effect of alkali metals are still ambiguous, researchers agree that alkali metal ions influence the rate of formation of iron carbides, which are the active phase in FTS [3].

Arslanfar *et al*. promoted Fe-Co-Mn catalyst with different alkali metals and ruthenium to observe their respective performances. They found that unpromoted catalyst yielded higher CO conversion than Cs and K promoted catalysts, but lower compared to Li, Rb, and Ru promoted catalysts. However, methane formation rate was much lower than alkali metal promoted catalysts. Consequently, they failed to present a catalyst with superior FTS performance. Yet, they still managed to improve their base catalyst by impregnating 10% MgO [29].

Nickel is another active metal for FTS, yet it is inapplicable for commercial processes due to its instability at real FTS conditions. Li *et al*. promoted unsupported Fe-based catalyst in the presence of Ni with different Fe/Ni atomic ratios. They suggested that at Fe:Ni=100:15 ratio, the catalytic activity is significantly enhanced, and more stable over 137 hours on

stream. However even at Fe:Ni=100:5 ratio, product selectivity shifts towards methane, thus lower FTS catalytic performance [54].

Feyzi *et al*. also studied the effects of amount of Ni promoted Fe loading on Al_2O_3 with a fixed Fe:Ni=2:3 molar ratio. They prepared bimetallic catalysts with a loading of 2-20 wt.% to test their performances at 360 °C and 1 bar. Optimum metal loading was detected as 16 wt.%, as CO conversion was maximized, and both CO_2 and CH_4 production rates were minimized. They also further promoted this optimal bimetallic catalyst with K_2S and $(NH_4)_2S$ separately to investigate any improvements in FTS performance. K_2S mildly increased CO conversion and significantly lowered CO_2 and CH_4 selectivity, whereas $(NH_4)_2S$ promoted Fe-Ni catalyst failed to provide an improved performance. Optimal K_2S loading on Fe-Ni catalyst was reported as 0.15 wt.% as it yielded maximum CO conversion along with minimum CO_2 and CH_4 selectivity. [55].

Alkaline-earth metals promote Fe catalyst similar to alkali metals, yet they are more stable at higher temperatures promoters because of their higher melting points. Li *et al*. promoted Si:Fe=5:100 base catalyst with 2 wt.% Mg, Ca, Ba, and Sr. Although CO conversion values were close to each other, Sr was designated as the best alkaline-earth metal promoter due to its high C_5^+ selectivity. Additionally, they determined the optimal Sr loading as 4 wt.% due to its improved stability [56].

Sulfur has been considered as a poisonous element for FTS catalysts and thus, there are limited research on sulfur promoted iron catalysts. However, small loadings of sulfur increases CO conversion and decreases methane formation [57]. Findings of Torres *et al*. indicate that 0.03 wt.% S on 5 wt.%Fe/α-Al_2O_3 improved per gram iron activity and decreased methane selectivity at 350 °C and 1 bar. They further impregnated S-Fe/α-Al_2O_3 with Na to improve FTS activity. Na-S-Fe/α-Al_2O_3 catalyst shifted product selectivity from methane to C_2-C_4 olefins and C_5^+ [23, 58].

Transition metals, such as Cu [21, 24, 33, 59], Cr [60, 61], Mn [62-65], Mo [66-68], Ta [60], V [60, 63], W [60], Zr [60, 62, 69] are other trending elements for Fe-based FTS catalyst promoters.

Similar to Fe-based catalysts, Co-based catalysts can be promoted with noble metals, transition metal oxides, and some rare earth metal oxides. However, unlike Fe promoters, main goal of Co promoters is to increase the reducibility of Co precursors to metallic Co agglomerates. Additionally, noble metal promoters are known to enhance Co dispersion over the support surface and increase stability of the catalyst [6].

Co-Ru bimetallic catalysts are very active and suitable for FTS due to efficient stand-alone performances of both metals. There are numerous studies to further improve their catalytic performances. Parnian *et al*. impregnated different concentrations of Ru on 15 wt.%/γ-Al_2O_3 to observe synergetic effect of two metals. They reported that CO conversion and C_5^+ selectivity is maximized with 0.2 wt.% Ru at 220 °C and 1 bar [70]. The same research group suggested that 0.3wt.%Ru-15wt.%Co/γ-Al_2O_3 yielded the best performance at the same reaction temperature and pressure, while halving the space time of the reactor [71]. Bimetallic Co-Ru catalysts on zeolites and CNTs also provided improved FTS activity [15, 72, 73].

Pirola *et al*. impregnated 0.1-2.0 wt.% Ru and Pt separately on 10wt.%Co/SiO_2 to test their FTS performances at 220 °C and 2.0 MPa. They found that 0.5wt.%Ru-10wt.%Co/SiO_2 gives the highest C_2^+ yield, and activity loss is significantly lower compared to other catalysts [74].

Alkali metals can be loaded with Co-based catalysts to shift the product distribution towards high alcohol chains at high pressures. Ishida *et al*. impregnated Na, Li, K, and Cs on 10wt.%Co/SiO_2 to investigate the change in formation rates of hydrocarbon and alcohol chains at 245 °C and 6.0 MPa. Their detailed analysis pointed out that Na is the most suitable candidate for high alcohol chain production due to its high selectivity to long chains, high conversion level, and short reaction time. The optimal amount of Na loading on a 10wt.%Co/SiO_2 was determined as 4 wt.%, since alcohol selectivity and CO conversion were almost constant above 4 wt.% loading [75].

Other transition metals Re [76, 77], Au [78], Zr [79], Mn [80], Ti [46], Zn, V, La [81], Ag [82], and Ni [73], and alkaline earth metals Mg [83], Ca [56, 84], and Sr [56] are investigated to improve Co-based catalysts.

High price of ruthenium precursors limits extensive usage of Ru-based catalysts. Ruthenium chloride is the cheapest and the most widely available precursor in the market. However, chlorine is known for its poisonous properties for CO and H_2 chemisorption. Nurunnabi *et al*. promoted Ru/Al_2O_3 and Ru/SiO_2 with Mn to increase FTS activity and catalyst stability. Mn loading was kept at one fourth of the support material. It was observed that Mn addition enhances FTS activity and increases resistance to deactivation on γ-Al_2O_3 by removing Cl atoms from the precursor by forming $MnCl_2$ on the surface and increasing metallic Ru density. Yet, Mn addition failed to improve performance on α-Al_2O_3 and SiO_2 supports [25].

2.1.4. Catalyst Preparation

Preparation of FTS catalysts conventionally involves 4 main steps: metal loading (if supported), drying, calcination, and reduction. Each step can be largely manipulated to optimize the size and morphology of the active phase.

There are numerous preparation methods applicable for FTS catalyst synthesis. Most widely used methods are precipitation (PR), wetness impregnation (WI), incipient to wetness impregnation (IWI), spray drying (SD), pore volume impregnation (PVI), slurry impregnation (SI), homogeneous precipitation (HP), evaporation-to-dryness, deposition precipitation, homogeneous deposition precipitation (HDP), pH precipitation (pHP), and microemulsion (ME).

Shimura *et al*. loaded 20 wt.% cobalt on Al_2O_3 support by wetness impregnation and homogeneous precipitation methods separately, and sequentially. XRD patterns and H_2-TPR results indicated that although HP yielded larger Co agglomerates and higher reducibility, WI method resulted in higher CO conversion and C_5^+ selectivity. They found that a two-step hybrid method, 10 wt.% Co by HP followed by 10 wt.% Co by WI gives the highest catalytic performance at 230 °C and 1.0 MPa [10].

Ma *et al*. investigated the effect of preparation methods for different cobalt loadings on Al_2O_3. XRD and results indicated that Co dispersion is higher when 15wt.%Co/Al_2O_3 is loaded by IWI method, and agglomerates are larger with SI. On the contrary for 25wt.%Co/Al_2O_3, SI yielded higher dispersion and smaller Co particles [79].

For promoted catalysts, order of impregnation of the metals has significant effect on the FTS performance. Miyazawa *et al*. observed the catalytic activity of sequentially impregnated, reverse sequentially impregnated, and co-impregnated Mn-Zr-Co/SiO_2 and reported that co-impregnation gives the highest C_5^+ yield [62]. Similarly, Parnian *et al*. suggested that highest chain growth probability and CO conversion is obtained when Co and Ru are co-impregnated on Al_2O_3 [71].

Calcination temperature, duration and medium composition substantially influence the catalytic surface such as surface area, pore volume, and pore diameter. Feyzi et al. investigated the effect of calcination temperature on unsupported 40wt.%Fe-60wt.%Ni between 500 °C and 700 °C. The catalyst calcined at 600 °C for 6 h yielded the best FTS performance regarding its high CO conversion and low CH_4 selectivity [42]. Same research group identified the optimum calcination conditions as 500 °C and 8 h for unsupported Fe-Mn-K catalyst [85].

Chu *et al*. calcined Co/Al_2O_3 and Co-Pt/Al_2O_3 at different temperatures ranging from 200 °C to 500 °C and tested at 190 °C and 1 bar. Although resultant conversion levels were similar due to low reaction temperature, calcination at 340 °C yielded the highest C_5^+ yield for Co/Al_2O_3, since higher temperatures hindered cobalt reduction. Co-Pt/Al_2O_3 calcined at 500 °C gave the best performance due to Pt promotion [86].

Fu *et al*. also investigated the effect of calcination temperature as well as medium composition on 20wt.%Co/CNT at 230 °C and 2.0 MPa. Thorough analysis revealed that calcination in argon medium at 200 °C is more active than that calcined in air. Additionally, they employed a novel technique, dielectric barrier discharge, and compared its performance with traditionally calcined catalyst. It was observed that dielectric barrier discharge plasma treated sample performed almost equally with the one calcined in Ar medium [87].

Reduction conditions significantly determine the chemical and physical properties of the active phase in a catalyst. Reduction temperature, duration, and medium composition are the parameters affecting the crystalline structure and catalytic activity in FTS.

Kang *et al*. evaluated the effect of reduction temperature on Ru-based catalysts and revealed that the size of Ru clusters increases as the temperature increase within 230 °C –

500 °C range [88]. This conclusion is in accordance with the results of Gonzalo-Chacón *et al*., which are studied under similar reduction conditions. Gonzalo-Chacón *et al*. also found that Ru-based catalyst yield similar CO conversion, but lower CH_4 selectivity when reduced at higher temperatures [89]. However for Co-based catalysts, high temperatures were found to enhance CO conversion, while also increasing CH_4 selectivity [50].

On the other hand, H_2 rich reduction medium may be redundant for Fe-based catalysts since the active phase is iron-carbides and iron-oxides, instead of metallic iron. Therefore these types of catalysts are usually activated under FTS conditions. Ding *et al*. activated Fe-Mn-K/SiO_2 under 265 °C, 0.5 MPa, and different syngas ratios, which are typical FTS conditions. They found that FTS activity is lowest when catalyst is activated under syngas ratio H_2:CO=4:1, and highest when the ratio is H_2:CO=1:4. In addition, catalytic activity loss is minimum within 316 h on stream [90].

2.1.5. Effect of Operating Conditions

In general, FTS can be classified in two groups according to their operating temperature: high temperature Fischer-Tropsch (HTFT) between 330 °C – 350 °C, and low temperature Fischer-Tropsch between 220 °C – 250 °C [6].

HTFT is usually carried out on Fe-based catalysts since iron-oxides favor high WGS kinetics on the active sites. This process is mainly developed to maximize gasoline and light olefin selectivity by using multi-tubular fixed-bed, fixed fluidized bed, or circulating fluidized bed reactors [91]. LTFT, on the other hand, can operate with iron, cobalt, and ruthenium based catalysts to produce long linear hydrocarbon chains such as waxes. Multi-tubular fixed-bed slurry-bed reactors are used to actuate LTFT [5].

There is a consensus in the literature on the effect of temperature in FTS reactions; it is widely accepted that increasing temperature enhances CO conversion both in HTFT and LTFT due to the exponential term of the Arrhenius Equation [92]. However, CH_4 selectivity is also enhanced due to high dissociation rate of the products from the active sites [17, 55, 70, 71]].

The primary function of operating pressure in FTS is tuning product selectivity. Selectivity to long hydrocarbon chains, and thus the chain growth probability, increases as the pressure increases between 0.1 – 2.0 MPa. Yet, the effect of total pressure on CO hydrogenation is still ambiguous.

Feyzi *et al*. conducted a series of studies on the effect of operating pressure for Fe-Ni-K/Al_2O_3 [55, 93] and Fe-Mn-K [85] catalysts. They detected a significant increase in C_5^+ selectivity with increase in pressure while CO conversion either remains unchanged [93] or slightly increases [55, 85]. Also, it is reported that CH_4 selectivity is slightly decreased in all cases. These findings are reinforced by the studies of Khodaei *et al*. [17], Mirzaei *et al*. [94], and Zhou *et al*. [95] which investigate Co-Mn/TiO_2, Co-Fe-K/SiO_2, and Co-Zr-SiO_2 catalysts, respectively. However, Almeida *et al*. reported 40.0 % activity loss with Co-Re/Al_2O_3 when the pressure is increased from 10 bar to 20 bar, while C_5^+ selectivity stayed almost the same [76].

Optimum H_2:CO molar ratio in the syngas for FTS is different for industrial LTFT and HTFT. Since WGS activity is negligibly low on Co and Ru based catalysts, LTFT processes require H_2 rich syngas, i.e. approximately 2.1. On the other hand, molar ratios between 0.7 and 1.0 are applicable for HTFT due to high WGS rates on Fe based catalysts [5, 13, 14]].

Parnian *et al*. extensively studied on the effect of syngas ratio with 15wt.%Co-(0.0-0.3)wt.%Ru/Al_2O_3. They reported that increasing syngas ratio from 1.0 to 3.0 significantly enhances CO conversion on the tested bimetallic catalysts at 220 °C and 1 bar. However, it also increases CH_4 formation and decreases C_5^+ selectivity [70]. A similar study of Xing *et al*. supported these results with Co/CNT under similar LTFT conditions. They attained a drastic increase in CO conversion while slightly decreasing C_5^+ selectivity [96].

Feyzi *et al*. performed LTFT with Fe-Ni-K/Al_2O_3 under different H_2:CO molar ratios changing from 1.0 to 3.0 at 230 °C – 270 °C and 1 bar. They found that a ratio of 2.0 gives the highest CO conversion and lowest CH_4 and CO_2 selectivity for all the tested temperatures [93]. The same group also tested a similar Fe based catalyst under HTFT conditions, 360 °C and 1 bar, with syngas ratios between 0.5 and 2.0. Their results indicate that 2.0 ratio gives

higher conversion and lower CH_4 selectivity at HTFT conditions as well, while the difference is rather mild [55].

FTS catalysts tend to lose activity as the time proceeds due to poisoning or coke formation on the active sites. Consequently, size of the active sites shrink and hence, favor shorter hydrocarbon chains and CH_4 selectivity [92].

Gas hourly space velocity (GHSV) is directly related with the time reactants spend in the reaction zone. Therefore, it has a massive effect on the conversion and product selectivity. Feyzi *et al*. investigated the effect of GHSV in Fe-Ni/Al2O3 at 260 °C and 1 bar between 1600 h^{-1} and 5600 h^{-1}. They results indicated a decrease in CO conversion as GHSV increases, as expected. However, lowest selectivity to CH_4 was attained at 2600 h^{-1}, and highest selectivity to C_5^+ products was obtained at 1600 h^{-1} [93]. The same catalyst was investigated under HTFT conditions by the same group between 2800 h^{-1} and 4000 h^{-1} at 360 °C and 1 bar. They found a similar decreasing CO conversion trend with increasing GHSV. Highest C_5^+ selectivity is obtained at 2800 h^{-1}, although the difference is slightly negligible [55].

2.1.6. Effect of Reactor Type

In industrial applications, HTFT process is conducted in multi-tubular fixed-bed, fixed fluidized bed and circulating fluidized bed reactors, whereas LTFT takes place in multi-tubular fixed bed and slurry-bed reactors [91].

Recently developing microchannel technology is considered as a suitable candidate for industrial FTS applications. Highly exothermic nature and wide product spectrum of FT reactions require advanced control systems to adjust catalytic activity and tune product selectivity. Microchannel reactors provide enhanced mass transfer properties within the channels, and intensified heat removal from the channels. Hence, they are good alternatives for energy efficient and highly selective industrial FTS applications [97].

Myrstad *et al*. [97] and Holmen *et al*. [98] compared FTS performances of fixed-bed reactor and microstructured reactor with Co-Re/Al_2O_3 in LTFT conditions. They maintained a significant improvement with CO conversion in microreactors, while product selectivity

was almost the same in both type of reactors. Holmen *et al*. also presented a detailed comparison between monolithic reactor and microreactor. They concluded that the former provides a higher rate of hydrocarbon formation and lower catalytic deactivation.

2.2. Data Mining Methods for Knowledge Extraction

Modern scientific studies and engineering applications are based on first principle models to describe physical phenomena in terms of mathematics. However, if the subject of interest is too complex to be modeled or the underlying first principles are yet to be known, data-driven models can provide hidden patterns, and hopefully useful knowledge. Recently developed and improved machine learning algorithms are the basis of knowledge extraction from the raw data in the tremendously large databases. The complete process transforming large amounts of raw data into meaningful predictive or descriptive models to discover knowledge is called *data mining* [99].

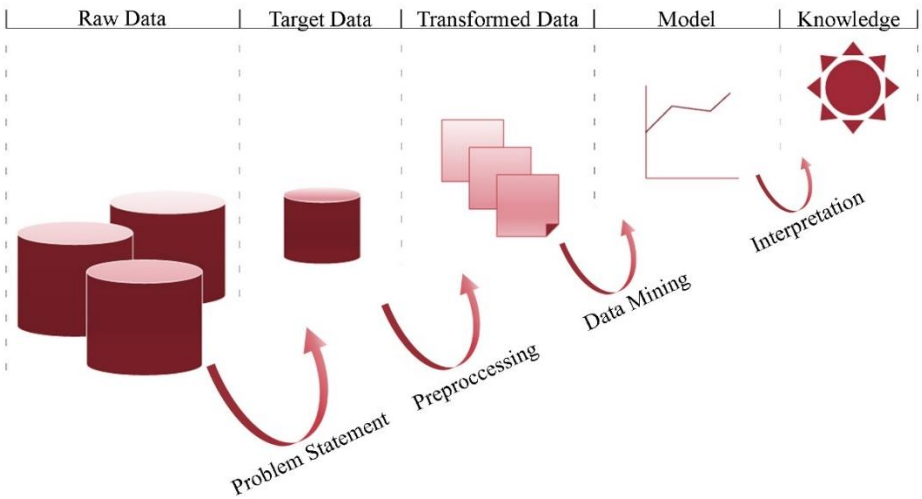

Figure 2.1. Data mining process.

A thorough data mining process involves a few steps as can be followed from Figure 2.1. The process is initialized by identifying a clear problem statement to address, and selecting appropriate dataset from the entire data presented in the world. This step includes choosing appropriate set of variables possibly related with the unknown dependency, and eliminating those unrelated with the target. A typical dataset is a very large matrix in which

the columns are called attributes, features or variables; and the rows are observations, or instances.

Since real-world data is imperfect, it may include outliers, missing, or even inexact information. Hence, the sieved database is preprocessed in the subsequent step for an optimal representation for the data mining algorithm. This step can involve outlier detection and removal, scaling to a predetermined range, dimensionality reduction, or dealing with the missing information.

Constructing a parsimonious model with an appropriate machine learning algorithm is the most essential step to develop structural patterns. K-nearest neighbor, artificial neural network (ANN), support vector machine, decision tree, random forest are a few of the fundamental supervised machine learning algorithms. Choosing the best model constructed by different algorithms is an additional challenge due to variety of indicators of a good model. Apart from the numerical indicators such as accuracy, error, and cost sensitive analysis; simplicity, scalability, and interpretability of a model has utmost importance [7].

The process is finalized by extracting interesting and useful patterns presented by the model. The results are carefully interpreted and conclusions are drawn accordingly. Interpreting the model and the results accurately is as challenging task as constructing an accurate model. Hence the model should generalize the complete domain of interest to make accurate predictions on future data, while being simple enough to interpret [99, 100].

Machine learning algorithms can be used for many tasks such as classification of categorical and ordinal attributes or prediction of numerical target variables through regression, revealing hidden structures and patterns, and association.

A typical database may comprise categorical target features such as “good” and “bad”, “black” and “white”, or ordinal variables such as “elementary school”, “high school”, and “college”. Alternatively, the target feature may be a numerical value such as age, price or height of an object. Previously mentioned supervised machine learning algorithms can be used to build a model based on the historical data, and predict future observations [7].

On the other hand, unsupervised learning algorithms such as Kohonen neural networks, hierarchical clustering, K-means clustering and principle component analysis are useful tools to group similar observations and reveal patterns between them [7].

2.2.1. Multiple Linear Regression

Multiple linear regression technique is one of the most widely used learning algorithms due to its simplicity and interpretability. It establishes a linear relationship between the input variables and the dependent continuous target variable. This relationship can be shown as in Equation 2.4 [101].

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon \tag{2.4}$$

Where, Y is the target variable, X_i is the input variable, β_j is the coefficient of multiple regression, p is the number of input variables, and ε is the residual. This equation can be elaborated in matrix form as in Equation 2.5.

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \cdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{pmatrix} \tag{2.5}$$

The coefficients are predicted by finding the least squares of the coefficient vector, i.e. minimizing the following summation.

$$\sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{p} x_{ij} \hat{\beta}_j \right)^2 \tag{2.6}$$

Therefore to be more concise, the following least squares equation yields the coefficient estimates.

$$\hat{\beta} = (X^t X)^{-1} X^t Y \tag{2.7}$$

Where, $\hat{\beta}$ is the vector of coefficient estimates, X^t is the transpose matrix of the input variables. Hence, the multiple linear regression model constructed as in Equation 2.8.

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p \tag{2.8}$$

2.2.2. Artificial Neural Network

Artificial neural network (ANN) is a powerful learning algorithm inspired by the biological neural networks in human brain. Learning capability and information processing capacity of human brain overwhelms the most complicated computer algorithms. It consists of nearly 10^{11} processing units, i.e. neurons, each having connections to 10^4 other neurons. All connections operate in parallel, giving its incredible computational power to undertake complex tasks such as image recognition and decision making. Hence, the idea behind constructing ANN is mimicing the nature of brain, and implement on computer algorithms [7].

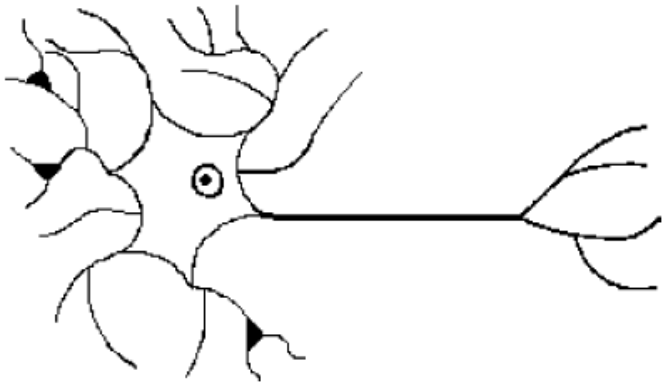

Figure 2.2. Sketch of a biological neuron [102].

A biological neuron (Figure 2.2) is the simplest structure of a network. It comprises numerous dendrites, and a single axon, all in contact with other neurons via contact points, called synapses. The nucleus of a neuron receives the stimuli from the presynaptic neurons via dendrites, generates a response if the sum of inputs is greater than a certain threshold, and passes the response along the axon to postsynaptic neurons.

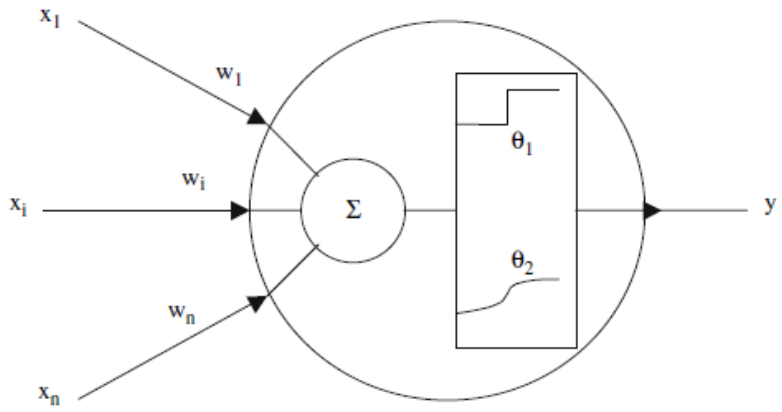

Figure 2.3. A typical artificial neuron model [102].

An artificial neuron is designed to adopt the same response generating process with a biological neuron, as shown in Figure 2.3. The neuron, represented by the wide circle, receives signals from surrounding neurons, x_i, via connections (dendrites) each having weights of w_i. Incoming signals and weights are summed by dot product and compared with the bias, θ, to produce an output, y. Output of a single neuron is scaled by an output activation function, f. Accordingly, the output of a neural network can be described as in Equation 2.9 [102].

$$\hat{y} = f\left(\sum_{i=1}^{n} w_i x_i - \theta\right) \tag{2.9}$$

Where, n is the dimension of the input variables. Llogistic sigmoid function is commonly employed as the activation function for neurons, as described in Equation 2.10.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.10}$$

Other widely used transfer functions are tangent sigmoid function, step function, and linear activation function, which are described below respectively [7, 101].

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2.11}$$

$$f(x) = \begin{cases} C_1 & if\ x > C_3 \\ C_2 & otherwise \end{cases} \tag{2.12}$$

$$f(x) = C_4 + C_5 x \tag{2.13}$$

Where C_1, C_2, C_3, C_4, and C_5 are constants. C_1 and C_3 usually equal to 1, and C_2 equals to 0, yielding the sign-activation function. In this specific case, the output of the transfer function is unity if the summation of the net input to the neuron exceeds the threshold zero; and zero if the net input has a negative sign. The sigmoid logistic and tangent sigmoid transfer functions are non-linear and differentiable, enables separation of linearly inseparable clusters by using backpropagation algorithm [7].

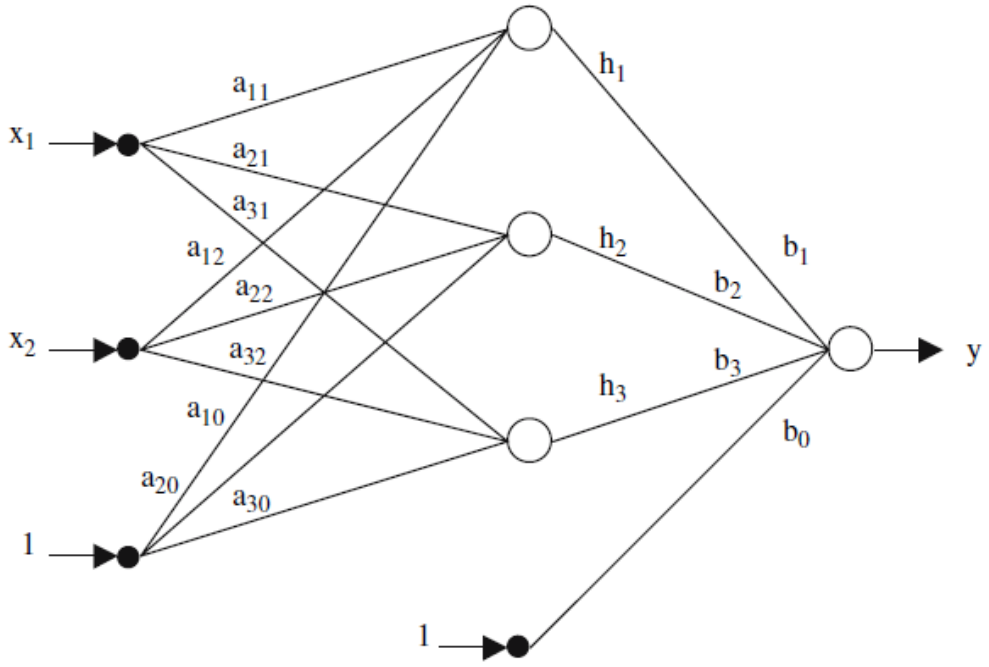

Figure 2.4. A scheme of feedforward network [102].

An ensemble of artificial neurons having properties discussed above constitute a neural network. Figure 2.4 demonstrates a two layer feedforward neural network topology with three neurons in a single hidden layer. The input layer, x_1 and x_2 are connected to the first hidden layer of neurons via connections with weights a_{ij}. An additional bias term is introduced via connections a_{10}, a_{20}, and a_{30}. The hidden neurons generate responses, h_1, h_2, and h_3, and transmit to output layer via b_1, b_2, and b_3. Another bias term is introduced to the

output neuron via b_0. The output neuron generates the final response, y, which is the desired output [102].

The weights of the connections are determined by implementing a backpropagation algorithm on the network topology. Unlike other supervised learning algorithms, training neural networks is usually carried out by an online learning technique. Instead of giving the whole sample space to the network at once, the instances are introduced one by one, allowing the network to update connection weights in every iteration. Complete diagram of neural network training process by backpropagation algorithm is demonstrated in Figure 2.5 [100].

Training the network starts with randomly initializing all the weights of the connections. Starting from the first hidden layer, the inputs are propagated forward by applying Equation 2.9 and one of the given activation functions on each neuron in the layer. After reaching the output layer, the error between the estimated output of the neuron and the known target value is propagated backward by updating the weights and biases. The error is calculated by Equation 2.14 when logistic sigmoid function is used as activation [100].

$$Err_j = \hat{y}_j(1 - \hat{y}_j)(y_j - \hat{y}_j) \tag{2.14}$$

Where $\hat{y}_j$ is the predicted output of the neuron j, and y_j is the known target value of the same unit. $\hat{y}_j(1 - \hat{y}_j)$ is the first derivative of the logistic function, therefore that part changes if any other differentiable activation function is used.

The error of a complete hidden layer is computed considering the weights of the connections between the units, as shown in Equation 2.15 [100].

$$Err_j = \hat{y}_j(1 - \hat{y}_j)\sum_k Err_k w_{jk} \tag{2.15}$$

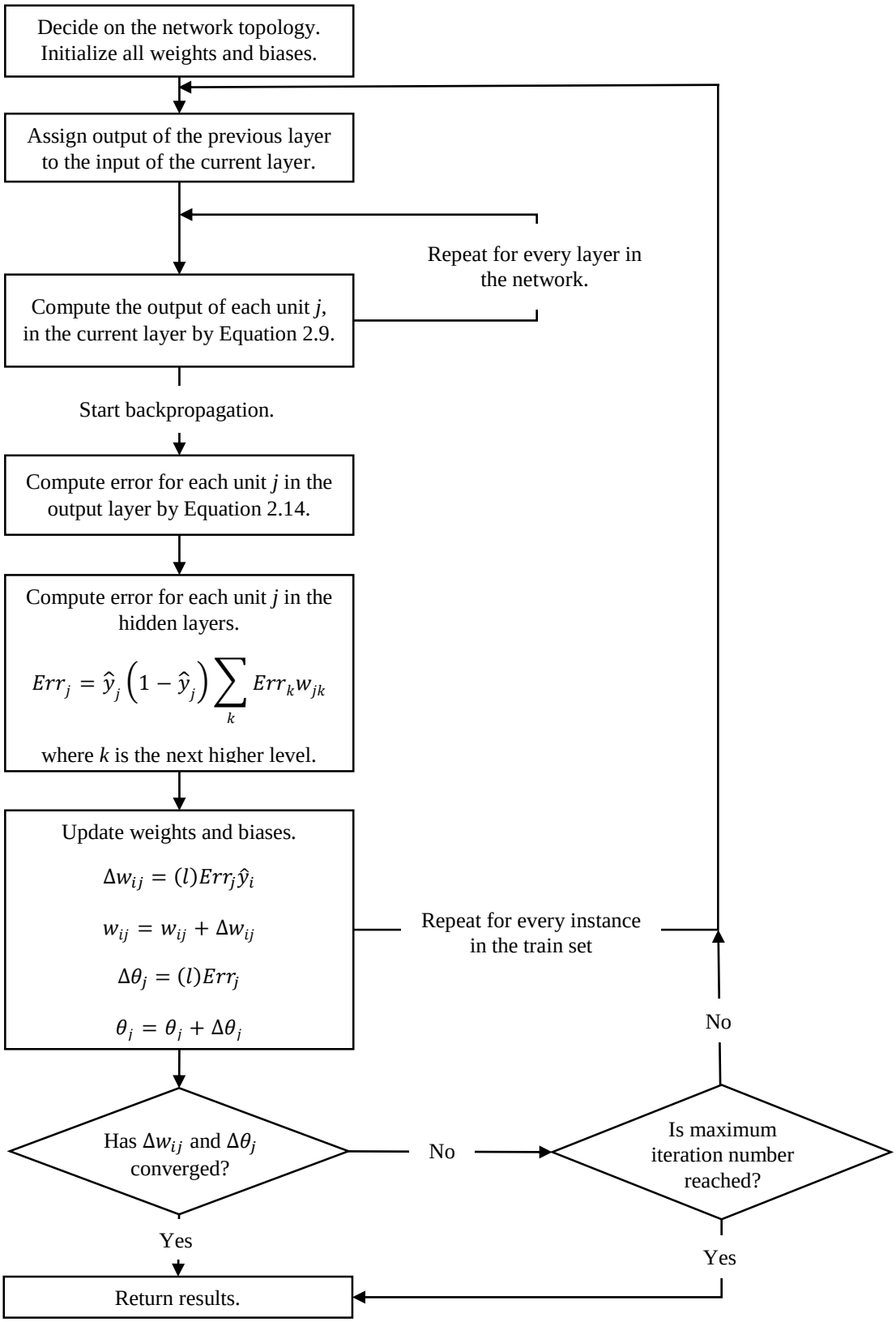

Figure 2.5. Neural network training process.

Where, w_{jk} is the weight of the connection from unit *j* to a unit *k* in the next higher layer, and Err_k is the error of unit *k*. Subsequently, the weights of the connections biases are updated by the following equations [100].

$$\Delta w_{ij} = (l) Err_j \hat{y}_i \tag{2.16}$$

$$w_{ij} = w_{ij} + \Delta w_{ij} \tag{2.17}$$

$$\Delta \theta_j = (l) Err_j \tag{2.18}$$

$$\theta_j = \theta_j + \Delta \theta_j \tag{2.19}$$

Where, Δw_{ij} is the change in the weight between the unit in previous layer and the current unit, $\Delta \theta_j$ is the change in bias of the current unit, and *l* is the learning rate. Learning rate is a numerical value typically between 0.0 and 1.0. If the learning rate is too small, the weights may stuck at a local minimum and increase computational time. On the other hand, too large a learning rate may oscillate the weights between two inadequate solutions, and never reach the global minimum. Backpropagation algorithm is mostly used with a gradient descent method to update the weights and biases, as given in Equations 2.14 – 2.19, but more complicated and much faster algorithms, such as Levenberg-Marquardt Method can also be implemented [100].

Neural networks have been used in various fields of physics, statistics, psychology, cognitive science, neuroscience, linguistics, and engineering [7]. The apparent learning power of these networks lies in its capability of handling highly non-linear input space, adaptability to nonstationary environment due to online learning mechanism, robustness even with missing or erroneous data [99]. These networks are also capable of handling other tasks than classification and regression; such as dimensionality reduction and unsupervised learning [7]. Nevertheless, an important drawback with these networks is that they are uninterpretable because of their working mechanism. They operate like black boxes, implying that none of the artificial neurons, or weights of connections have a physical comprehensible meaning [100].

2.2.3. Decision Tree

A decision tree is a supervised model, recursively splitting the complete heterogeneous database into more pure clusters by using logical arguments as rules. As demonstrated in Figure 2.6, a tree comprises one root (starting) node and several internal nodes, where an attribute is tested based on a decision criterion. Each internal node divides the entering input set into two or more subsets, and transmits them to the next internal node, or terminal (leaf) nodes via branches. Terminal nodes yield the prediction of the including observations [7].

Decision trees can handle numerical (continuous) variables as well as categorical (discrete) ones and thus, they can be adopted both for classification and regression type problems [100]. Figure 2.6 demonstrates one root node and two internal nodes partitioning and labeling the dataset into five subsets using both continuous and discrete variables. The shaded part in each node represents the fraction of "Class 1", and remaining part is "Class 2". The root node in the figure is highly heterogeneous, containing both classes almost equally. The purity of each class increases as we follow the branches until the terminal node.

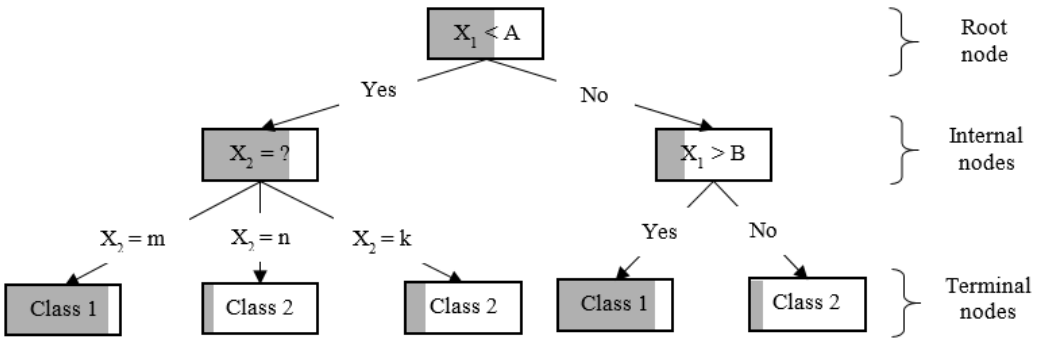

Figure 2.6. A simple decision tree.

A decision tree is mostly trained by a top-down recursive divide-and-conquer technique. CART and C4.5 are two benchmark algorithms on decision tree induction, and they follow a similar strategy, described in Figure 2.7 [100].

While growing (training) a tree, there are three frequently used attribute selection measures for the splitting criteria in the internal nodes and the root node, namely,

information gain, gain ratio, and Gini index. Choosing the best splitting measure requires a trial-and-error process, since none of them are unconditionally superior to others [100].

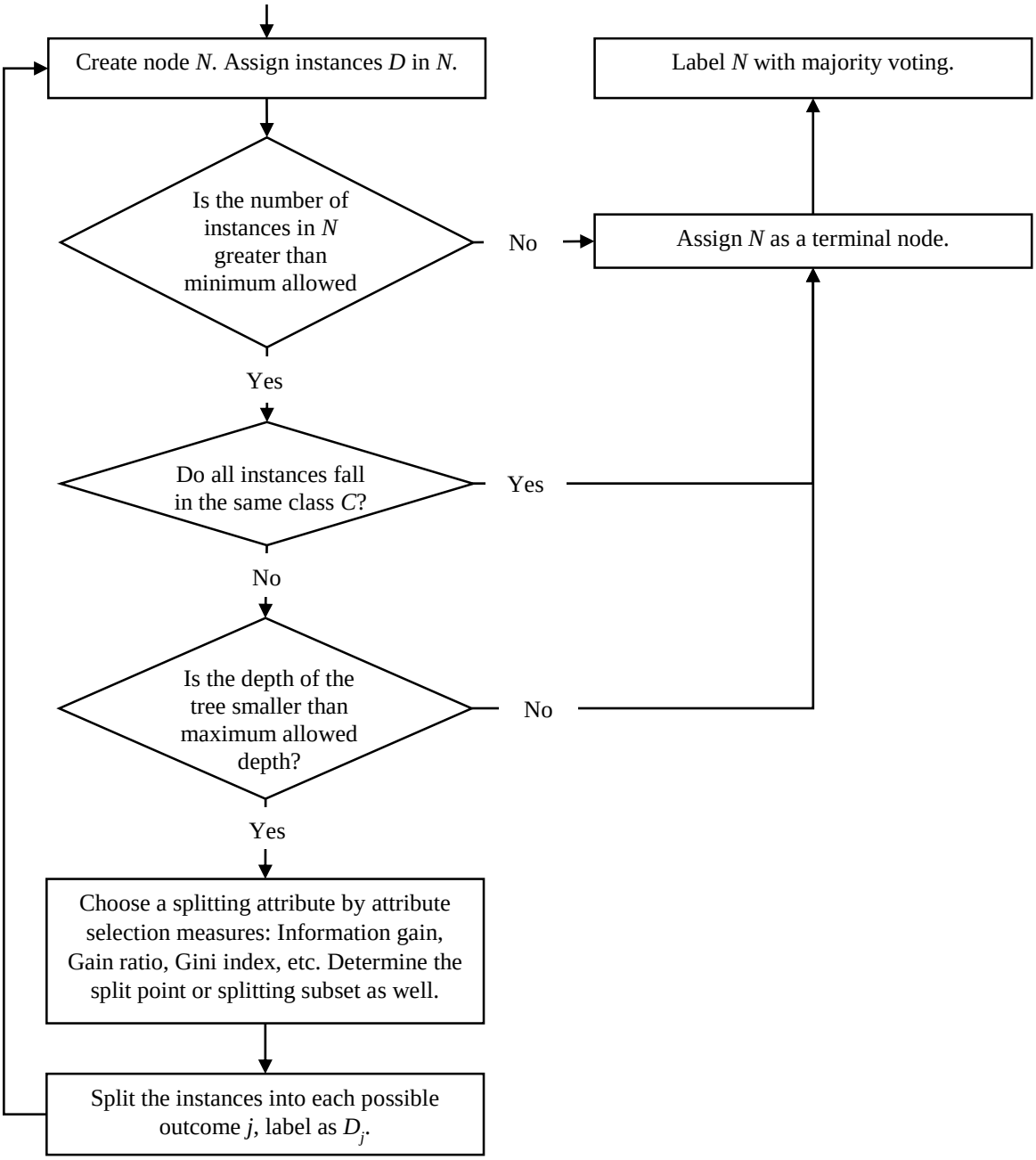

Figure 2.7. Basic algorithm to grow a tree.

In information gain method, the attribute with the highest information gain is chosen to split the attribute in the given node. The expected information or entropy, in this sense, is defined as in Equation 2.20 [100].

$$Info(D) = -\sum_{i=1}^{m} \frac{C_{i,D}}{D} \log_2 \frac{C_{i,D}}{D} \tag{2.20}$$

Where, C_i is the number of instances of the ith class in the current node, D is the total data partition in the same node, and m is the number of classes in the target variable. The information required to obtain a pure classification in the node is given in Equation 2.21 [100].

$$Info_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \times Info(D_j) \tag{2.21}$$

Where A is an arbitrary attribute, v is the number of possible values that A can have, and D_j is the partition after splitting. Since the aim of partitioning is to acquire subsets as pure as possible, $Info_A(D)$ value is expected to be small for a good split.

The difference between the original information and the information required after splitting gives the information gain, as shown in Equation 2.22 [100].

$$Gain(A) = Info(D) - Info_A(D) \tag{2.22}$$

Therefore, $Gain(A)$ gives the increase in information by branching the attribute A. The attribute giving the highest gain is assigned as the splitting attribute in the node [100].

Information gain approach is used in ID3 algorithm, one of the first famous decision tree induction algorithms. C4.5, a successor of ID3, employs gain ratio to select the splitting attribute, whereas CART (Classification and Regression Tree) utilizes Gini index for the same purpose.

Popularity of decision trees stems from its lower training cost in terms of computational time compared to other machine learning algorithms. Additionally, it is a non-parametric algorithm and requires almost no a priori knowledge on the domain. The most important leverage of decision trees is that it is intuitive and easily interpreted by the user [100].

2.2.4. Random Forest

Random forest in an ensemble method, in which more than one decision tree are grown. Each tree is trained by randomly and independently selected attributes, instead of all the attributes in the dataset. In addition to random selection of the attributes, the instances used for growing a tree are selected by bootstrapping [100].

The procedure to plant a forest with *k* trees starts with sampling the instances, *D*, with replacement. Thus, some of the instances may occur more than once while others may be excluded from the sample space. Subsequently, the number of attributes in a tree, *F*, is selected such that *F* is much smaller than total number of attributes in *D*. The CART algorithm may be employed to grow the trees individually. When a test instance is applied on the forest, each tree votes for the winning class, and the instance is assigned by majority voting [100].

Random forests are comparably accurate while being robust to outliers and errors in the dataset. Generalization ability of a forest converges as the forest size increases and thus, eliminates the problem of overfitting. Since random forests are modeled with considerably fewer attributes in each tree, they work very efficiently compared to other ensemble methods [100].

2.2.5. Principal Component Analysis

Principal Component Analysis (PCA) is mostly used for dimensionality reduction in databases with large number of attributes. PCA aims to explain the most variance of the original dataset by reducing the dimensionality of the column space. Since initial data directions are preserved as much as possible, this algorithm is used to visually describe large datasets on 2-D or 3-D plots [103].

The basic procedure for PCA starts with normalizing the input columns to prevent dominance of too large attributes over too small ones. *n* orthonormal vectors (principal components) are computed such that they cover total variance of the normalized input data. The original input data can be represented as linear combinations of the principal components. The principle components are sorted in order of decreasing variance they

individually explain, and *k* of them are selected to represent the input data in the new reduced column spaces. *k* value is usually selected such that the loss in variance of the input is minimum while reduction in dimensionality is maximum [100]. Yet, it can be chosen to be 2 or 3, if the input space is to be visualized on a plot.

2.3. Data Mining on FTS and Heterogeneous Catalysis

Data mining algorithms are widely used in diverse fields and purposes. In banking, credit applications, fraud detection, and stock markets are modeled by using these algorithms. Engineers benefit from such algorithms for optimization, control, and troubleshooting. Medicine, telecommunications, physics, astronomy, and biology are other fields that deal with large amounts of data and use data mining algorithms to reveal a pattern or for predictive purposes [7]. Recently, it has been used for identifying protein structures and functions [104-106], modeling biomass growth [107], designing industrial processing units [108, 109]. Application of data mining algorithms on heterogeneous catalysis field started with artificial neural networks by Hattori *et al.* [110].

The first study on FTS integrated with a data mining algorithm was published by Sharma *et al*. in 1998 [111]. It was reported that ANN can be used to optimize C_5^+ liquid hydrocarbon selectivity by predicting $CO+H_2$ conversion and steady state product concentrations under different reactor operating conditions. Error-back-propagation algorithm was used to train the multi-layer feed-forward network. 5 separate and independent networks were constructed with 3 inputs (reactor pressure, space velocity, and reaction temperature) and 1 output neuron structure to model 5 different quantities, i.e. $CO+H_2$ conversion, C_5^+ liquid hydrocarbon concentration, methane concentration, oxygenate concentration, and hydrocarbon concentration. 7 experimental runs were used to train the network model, and 2 experimental runs were used for the test set. Root mean square errors for the train set and test set were reported between 0.05 – 0.19 and 0.06 – 0.17% respectively for 5 different networks, which are quite satisfactory and promising for future studies.

It has been shown that different data mining algorithms can be integrated to build improved prediction models. Rahimpour *et al*. developed the accuracy of ANN models on FTS performance by combining with genetic algorithm (GA) [112]. Effects of resident time,

pressure and temperature on molar percentage of CH_4, CO_2 and CO were investigated through 6 dataset including 45 data. Mean square errors were reported between 0.023 and 0.052 with ANN model, and 0.002 and 0.005 when GA-ANN model was used.

Studies covering the entire literature in the field of heterogeneous catalysis are very limited in number [9, 113-115]. It was reported that a comprehensive and accurate statistical model can be constructed to infer useful trends, rules and correlations between the performance of the catalyst and interaction between catalyst preparation and operational conditions.

Zavyalova *et al*. studied on oxidative coupling of methane through 1870 datasets from ca. 420 full-text references on catalyst compositions and their performances. ANOVA, correlation analysis, and decision tree algorithms were implemented to make a descriptive interpretation. The study reports that 18 catalytic elements with their binary and ternary interactions have significant effects on the catalytic activity and C_2 selectivity [113].

Gunay and Yildirim studied on CO oxidation over Cu-based catalysts by constructing an ANN model with 1337 data points from 20 research publications in the literature. CO conversion was successfully linked to the catalyst preparation variables and operating conditions [114].

Apart from performance studies, global reaction rate models were constructed by ANN algorithm using past data in the literature. Gunay and Yildirim studied CO oxidation kinetics over Au based catalysts by using 882 data points extracted from 19 research papers. The resultant power law type global reaction rate model gave acceptable predictions for the experimental results [9].

Odabasi *et al*. studied water gas shift reaction over noble metal catalysts (Pt and Au) through 4360 experimental data points extracted from 84 research publications. Decision tree algorithm was implemented to determine empirical relations between the catalyst preparation variables and CO conversion. ANN was used to determine their relative effects on the output. Lastly, support vector machine was used to predict CO conversion of unstudied experimental conditions [115]. All three algorithms were successfully applied to

water gas shift over noble metal catalyst systems, implying that hidden rules and trends in such recondite catalytic systems can be revealed by data mining algorithms.

3. COMPUTATIONAL DETAILS

In this section, construction steps and the details of the FTS performance database was thoroughly explained first. Subsequently, the methodology followed to extract knowledge from this database, as well as application of the machine learning algorithms used were described.

3.1. Experimental Data Collection

The first step of this work was constructing a FTS performance database considering retrospective articles in the field. Tremendous amount of studies on FTS have accumulated over years, and it keeps accumulating with an increasing trend. Figure 3.1 shows the number of annual publications between 2005 and 2014, according to Web of Science™. It can be observed that the interest in FTS has been doubled in the last 10 years.

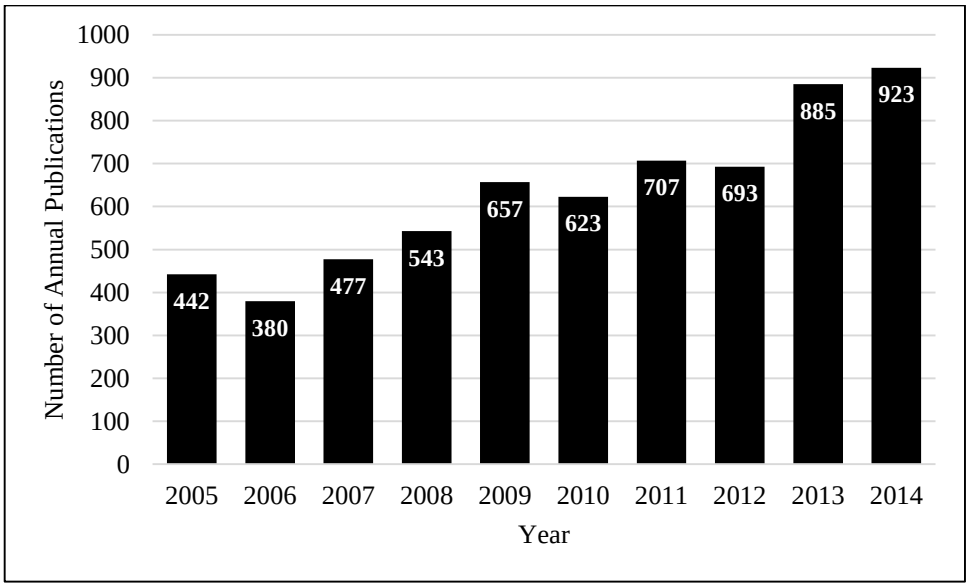

Figure 3.1. Number of annual publications on FTS between 2005 and 2014.

In this study, 371 research articles were reviewed to construct the FTS performance database, and 171 of them [10, 13, 15, 17-21, 23-28, 30-44, 47, 50-52, 54, 56, 58, 59, 62-70, 73, 74, 76-78, 80, 82-86, 93-97, 116-222] were designated to be suitable for our purpose. A "suitable research article", in this case, refers to those reported the input and output variables

with a clear statement of the quantity, quality, and the procedure without any obscurity. 1361 unique experimental runs were extracted from the selected articles, which yielded 4098 data points in total.

Each row was represented in a 72 dimensional space with 10 attributes, which were demonstrated in Table 3.1.

Table 3.1. Input attributes and their identities.

Attribute	Identity
Base metal	Co, Fe, Ru
Promoter metal	Ag, Au, Ca, Cu, K, La, Mg, Mn, Mo, Na, Ni, P, Pt, Re, S, V, Zr
Support	CeO2, CNF, H-ZSM-5, MWCNT, OMC, SBA-15, SiC, SiO2, SWCNT, TiO2, ZnO, ZrO2, ZSM-5, α- Al_2O_3, γ-Al_2O_3
Base metal loading method	Precipitation, wetness impregnation, incipient to wetness impregnation, spray-dried, slurry impregnation, homogeneous deposition precipitation, pH precipitation, others
Calcination conditions	Calcination temperature, calcination time
Reduction conditions	Reduction temperature, reduction time, and reduction medium
Promoter loading method and order	Sequel impregnation, coimpregnation, reverse sequel impregnation, coprecipitation
Physical properties	Surface area, pore volume, pore diameter, metal particle size
Operating conditions	Time on stream, temperature, pressure, GHSV, feed composition
Reactor type	Fixed bed, CSTR, Microchannel

Special care was taken while selecting these attributes to avoid inclusion of rarely used elements/methods/properties since they will not lead to globally generalized and robust models. Only the attributes appeared at least in five experiments were included in the database.

Active metal is the most significant characteristic of a FTS catalyst, since it directly influences the design of the reactor and the operating conditions in an industrial process. In this work, the database included three active metals (Co, Fe, and Ru) as numerical attributes. Although Ni is also active for FTS reactions, it is highly unstable and thus, there are only few works on a standalone Ni catalyst; hence this metal did not taken into consideration. Table 3.2 shows the number of data points with each base metal with their corresponding ranges in weight per cent (wt.%). It should be noted that some of the research papers reported base metal loading in terms of mole fraction, which were converted to wt.% for the database. Also, a few of the studies took the weight of the support as basis, i.e. the ratio of the metal to support was reported. These results were also converted to the basis of complete catalyst weight.

Table 3.2. Number of data points with respect to base metals and their ranges.

Base Metal	**Number of Data Points**	**Range (wt.%)**
Co	2069	0 – 40.1
Fe	1900	0 – 100.0
Ru	686	0 – 10.0

Promoter loading in a FTS catalyst is another essential attribute affecting the catalytic activity and stability. There were 17 elements in the database used as promoters, namely Re, Ag, Au, Na, K, Mg, La, Ca, Cu, Mn, Mo, Ni, V, Zr, P, Pt, and S. The number of data points with these elements, and their corresponding loading ranges in wt.% were given in Table 3.3.

Table 3.3. Number of data points with respect to promoters and their ranges.

Promoter	Number of Data Points	Range (wt.%)
K	1137	0 – 10.0
Cu	742	0 – 95.8
Mn	637	0 – 74.7
P	380	0 – 51.4
Ni	139	0 – 100
Pt	119	0 – 5.0
Re	117	0 – 4.5
Ca	99	0 – 8.5
Zr	95	0 – 15.0
V	36	0 – 18.2
Mo	32	0 – 72.0
La	24	0 – 4.5
Na	21	0 – 3.8
Mg	18	0 – 0.8
Au	17	0 – 5.1
Ag	13	0 – 2.8
S	12	0 – 0.4

Support of a FTS catalyst influence dispersion of the active metal, and the size of the active site agglomerates by providing high surface area, pore diameter and pore volume. It can also chemically enhance the electronic state of the active sites. The dataset included 15 types of supports, namely γ-Al_2O_3, α- Al_2O_3, SiO_2, SiC, ZnO, ZrO_2, CeO_2, single-wall carbon nanotubes (SWCNT), multi-wall carbon nanotubes (MWCNT), carbon nanofibers (CNF), H-ZSM-5, ZSM-5, SBA-15, ordered mesoporous carbons (OMC), and TiO_2. Each type of support was represented by a numerical attribute, regarding the balancing weight per cent of base metal and promoter element. Table 3.4 shows the number of data points with each type of support and its corresponding weight per cent.

Table 3.4. Number of data points with respect to supports and their ranges.

Support	Number of Data Points	Range (wt.%)
γ-Al_2O_3	1524	0 – 98.0
SiO_2	1438	0 – 99.0
MWCNT	372	0 – 95.2
SBA-15	149	0 – 90.0
ZSM-5	144	0 – 88.0
TiO_2	90	0 – 94.5
α- Al_2O_3	75	0 – 95.0
SWCNT	66	0 – 90.4
H-ZSM-5	58	0 – 86.0
ZrO_2	42	0 – 90.0
ZnO	34	0 – 95.9
CeO_2	33	0 – 97.0
SiC	16	0 – 90.0
OMC	14	0 – 90.0
CNF	10	0 – 95.0

The FTS catalysts in the database were synthesized by seven different catalyst preparation methods, namely precipitation (PR), wetness impregnation (WI), slurry impregnation (SI), spray drying (SD), incipient to wetness impregnation (IWI), homogeneous deposition precipitation (HDP), pH precipitation (pHP), and others. “Others” included preparation techniques which were very few in number, such as pore volume impregnation (PVI), homogeneous precipitation (HP), evaporation-to-dryness (ED), deposition precipitation (DP), and microemulsion (ME). Binary representation was used for this attribute, since catalyst preparation method is a categorical attribute. Logical true, “1”, was used for the preparation method used in that instance, and logical false, “0”, was used for the others. Some of the experimental runs included catalysts prepared by a combination of two different methods. For those instances, both of the preparation methods were assigned

a logical true. Figure 3.2 shows the distribution of the catalyst preparation techniques with respect to the number of data points.

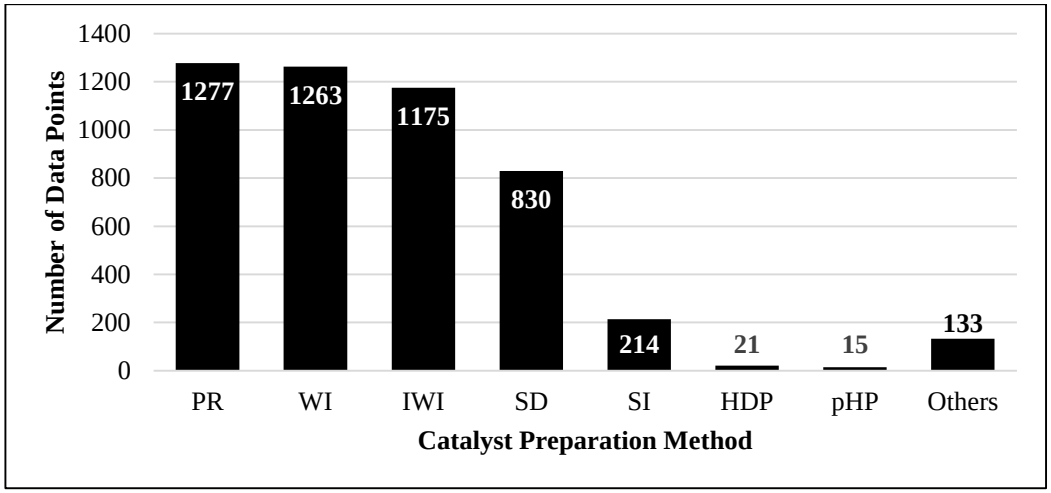

Figure 3.2. Distribution of the catalyst preparation techniques.

Most of the catalysts were calcined before running the experiment. Calcination temperature and calcination time were two continuous attributes related to the calcination conditions. The temperature values were given in Kelvin (K), and duration was given in hours (h). If there were multiple calcination steps at different temperatures, the one with the higher temperature was recorded. If an uncalcined catalyst was used in the experiment, calcination temperature was recorded as 298 K for convenience. Figures 3.3 and 3.4 show the distributions of the calcination temperature and calcination time with respect to the number of data points.

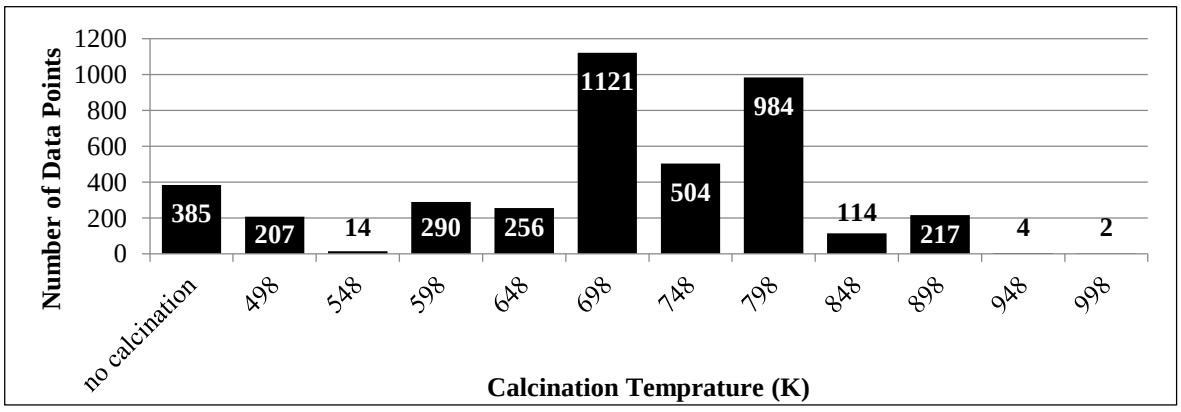

Figure 3.3. Distribution of the calcination temperature.

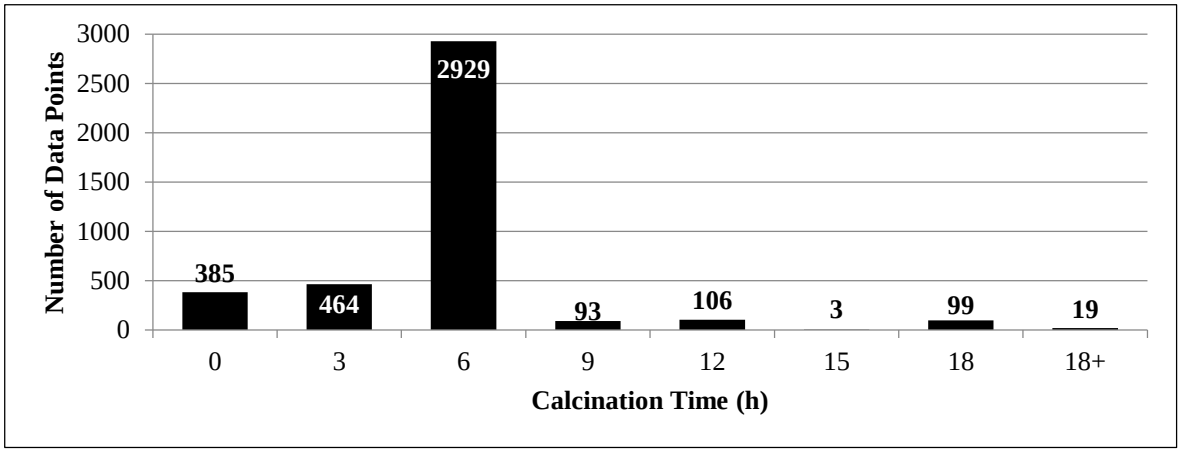

Figure 3.4. Distribution of the calcination time.

Reduction conditions determine the size and crystal structure of the metallic phases in the catalyst and thus, another attribute for the database. The reduction conditions comprised temperature, time, and composition of the gas used. The units in the research papers were converted into Kelvin (K), hour (h), and volumetric per cent (vol. %), before being recorded to the database. Similar to the calcination variables, if the unreduced catalyst was used in the experiment, temperature value was assigned 298 as K, and the reduction time and gas concentrations in the reduction medium were zero. Figures 3.5 and 3.6 demonstrate the distributions of reduction temperature and reduction time with respect to the number of data points.

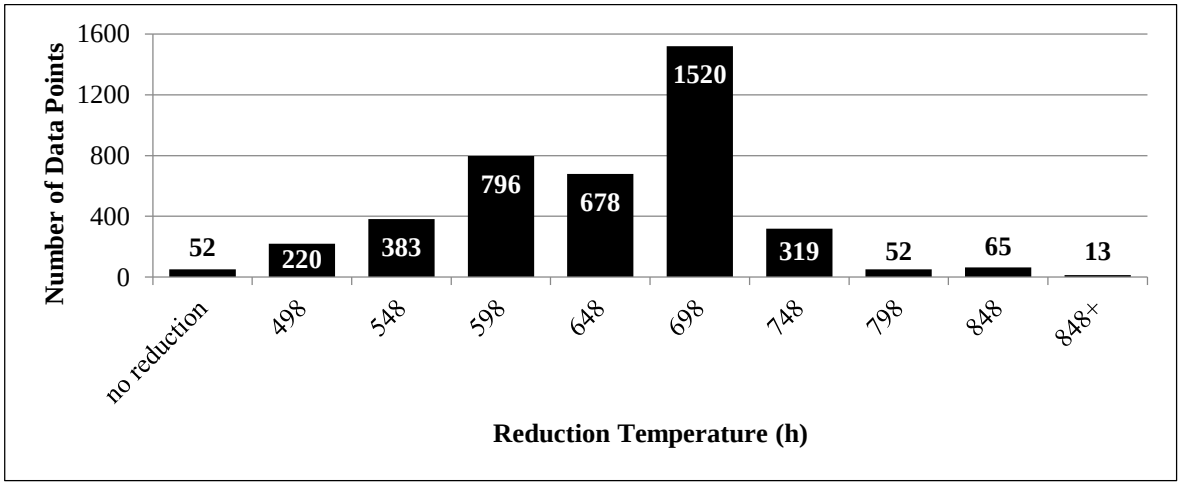

Figure 3.5. Distribution of the reduction temperature.

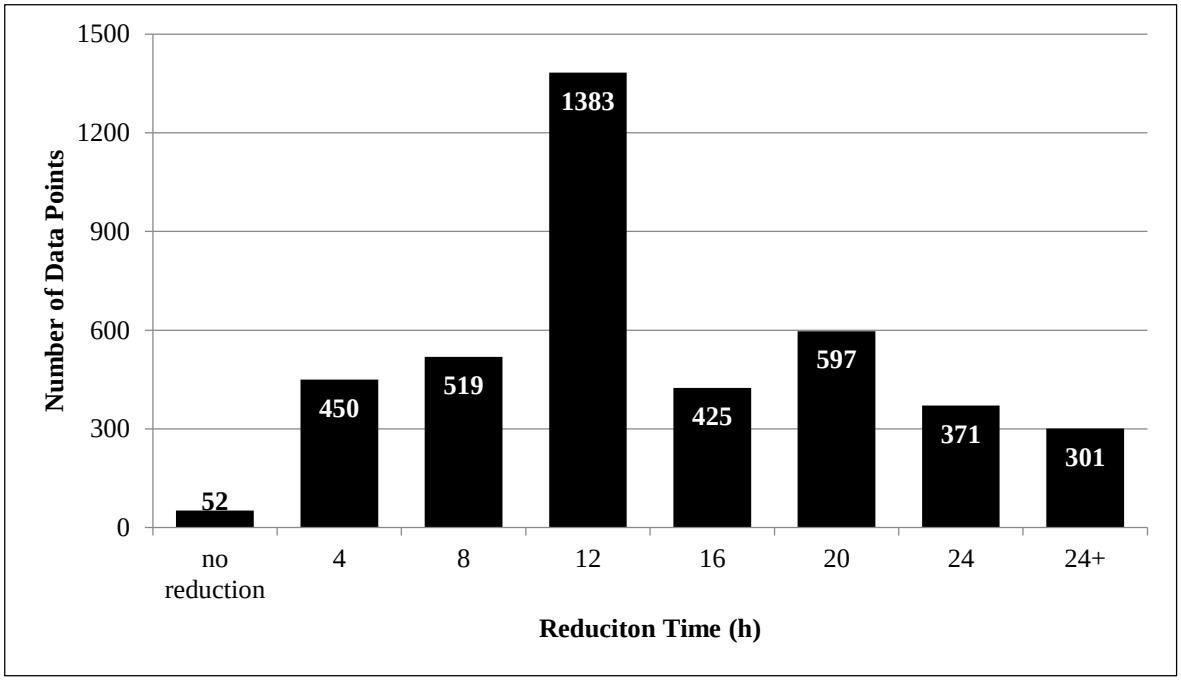

Figure 3.6. Distribution of the reduction time.

Most of the experimental runs had reduced the catalyst under pure hydrogen medium. However, some of the catalysts were reduced under hydrogen flow diluted with N_2, He, or Ar, all of which were classified as "Inert". Also, some catalysts were activated with syngas, under FTS reaction conditions, which required the "CO" attribute in the reduction medium. The ranges of medium concentrations of the attribute were provided in Table 3.5.

Table 3.5. Range of the reduction medium concentrations.

Medium	Range (vol. %)
H_2	0 – 100
CO	0 – 100
Inert	0 – 95.0

Promoter loading method and order has been another subject of investigation in the literature. Promoter metals in the database were either loaded by precipitation or impregnation. Addition by impregnation was divided into three categorical attributes with respect to the order of base metal loading and promoter loading. If the promoter was

impregnated after the base metal, “sequential impregnation” was assigned a logical true, and the opposite condition was marked as “reverse sequential impregnation”. If the promoter was simultaneously impregnated with the base metal, the case was denoted as “co-impregnation”. If more than one promoter was used to synthesize the catalyst, the method and order was recorded regarding the relatively important element. The distribution of promoter loading method and order was represented in Figure 3.7.

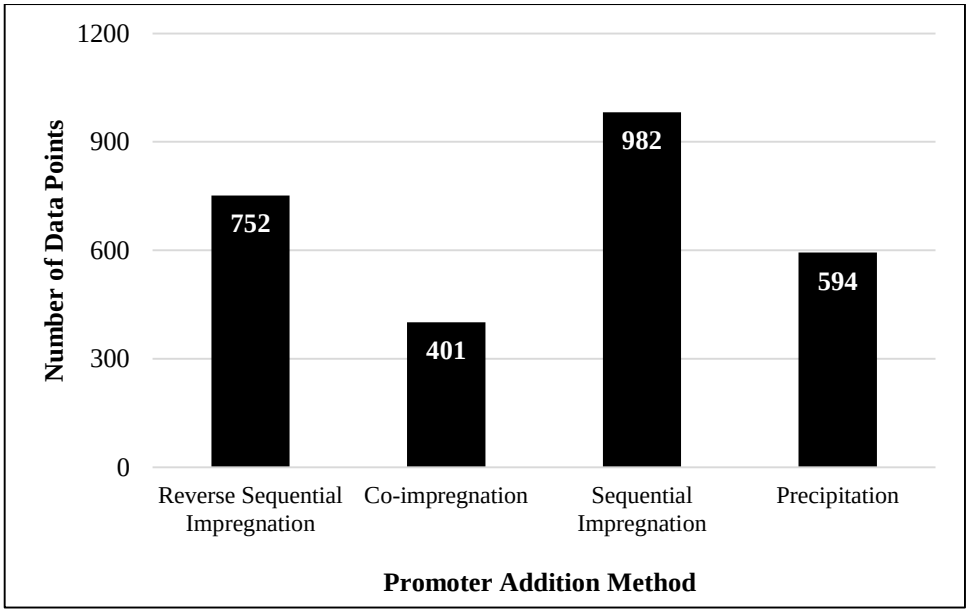

Figure 3.7. Distribution of the promoter addition methods.

The attributes related with the physical properties of the catalyst, namely surface area, pore volume, pore diameter, and metal particle size, are a special kind of exceptional variables that should be treated carefully. These properties are normally functions of the previously mentioned variables and including them in the dataset has a potential for causing multicollinearity. Yet, most of the research papers employ minor modifications on catalyst preparation steps to tailor the textural properties of the catalyst. Including all of these modifications would substantially increase the dimensionality and sparsity of the dataset. Therefore, it was assumed that including the physical properties of the catalyst instead would take the minor modifications into account. Besides, the correlation analysis was performed and it was found that only surface area and pore volume were mildly correlated with a correlation coefficient of 0.62, and the rest of the properties were either weakly correlated or non-correlated. The distribution of these variables were given in Figure 3.8.

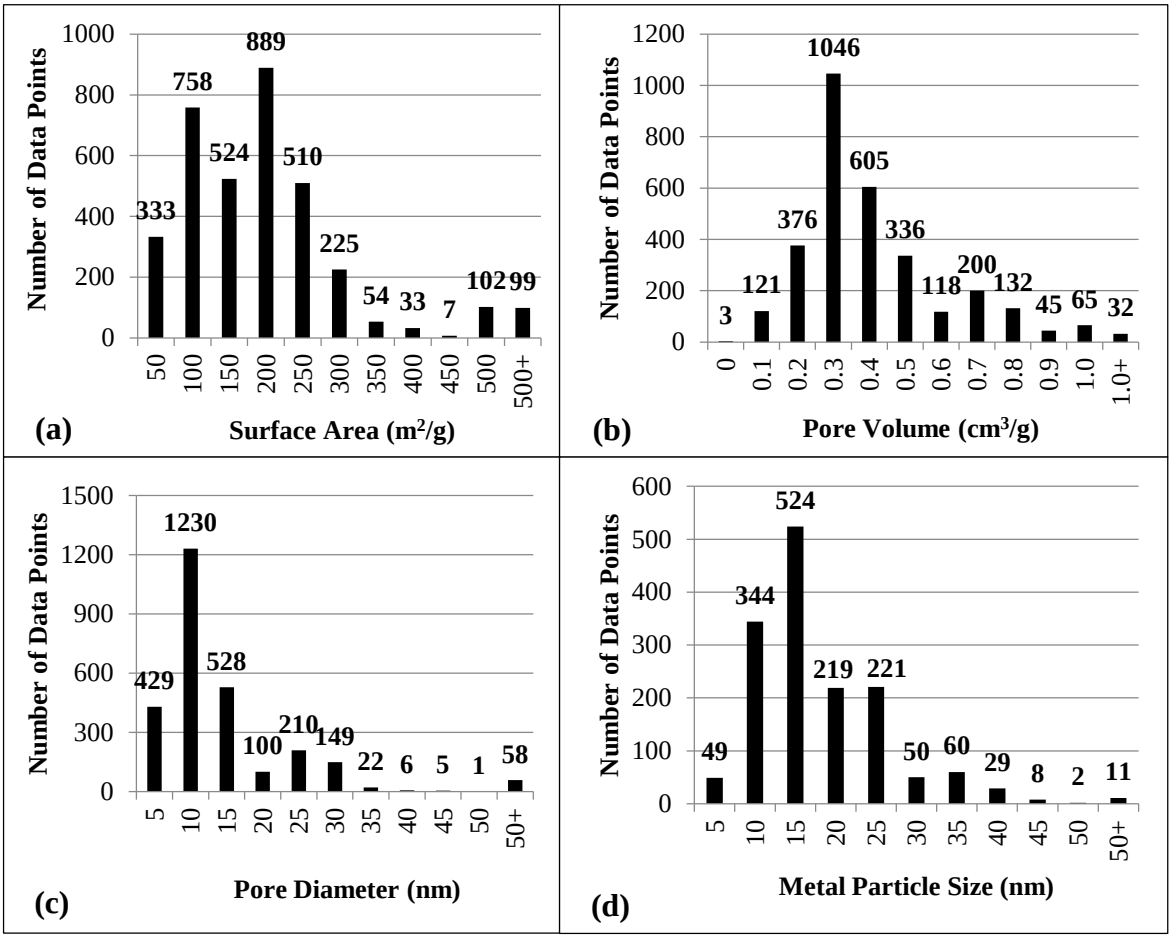

Figure 3.8. Distribution of the physical properties: (a) surface area, (b) pore volume, (c) pore diameter, (d) metal particle size.

Time on stream (TOS), operating temperature, operating pressure, gas hourly space velocity (GHSV), and feed composition were designated as continuous operational variables in the dataset. The units in the research papers were converted to h, K, bar, mL/g.min, and vol. %, respectively, to maintain consistency in the dataset. However, an assumption was required to be able to convert GHSV in h^{-1} to mL/g.min. The unit h^{-1} withholds any kind of information on the catalyst weight in the reactor. Therefore, it was assumed that the reactors in the experiments were packed with the bulk density of the support material. Accordingly, the GHSV in mL/g.min was calculated by dividing GHSV in h^{-1} to the commercial bulk density of the support. The distributions of the operating variables were given in Figure 3.9 and Table 3.6.

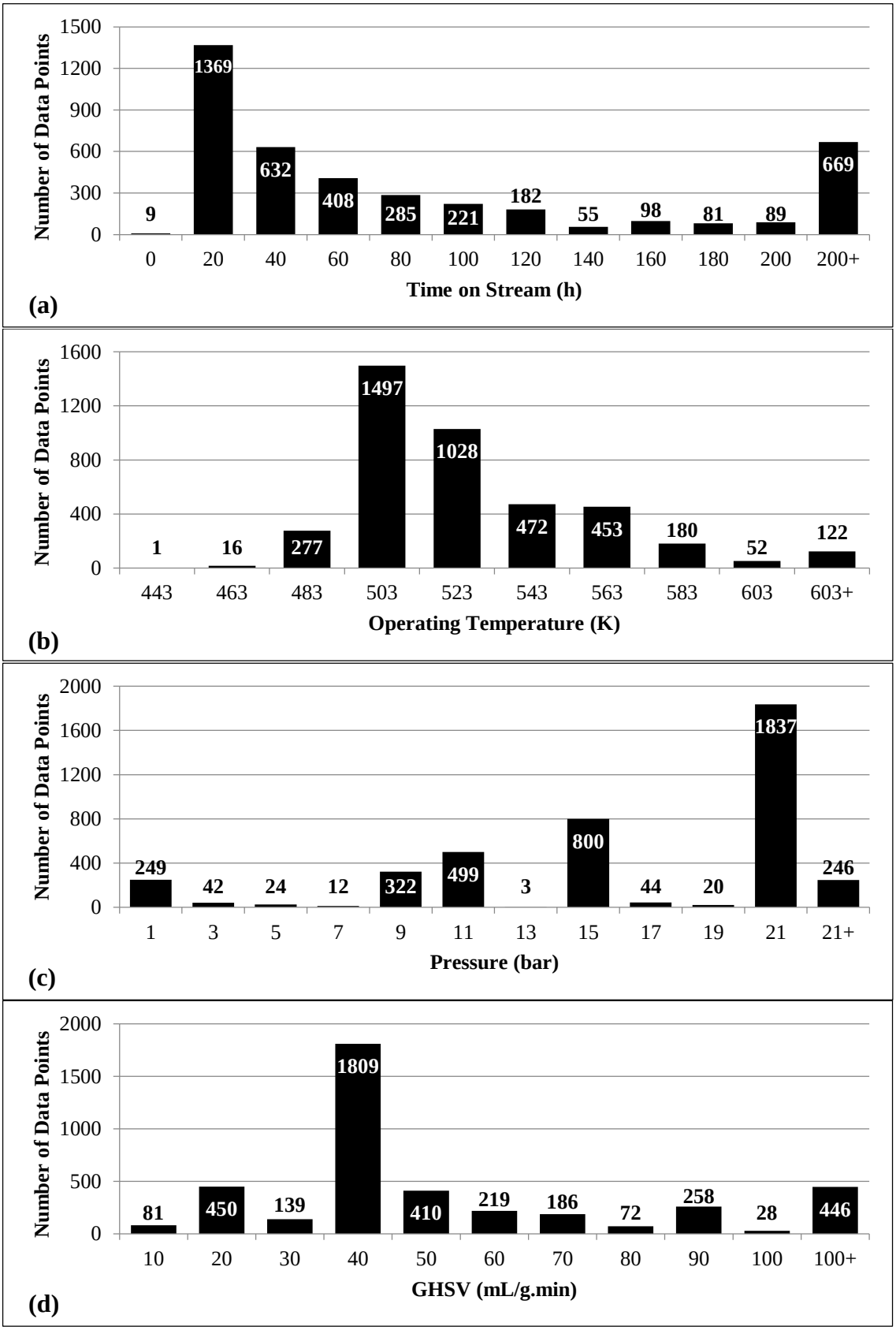

Figure 3.9. Distribution of the operating variables: (a) time on stream, (b) operating temperature, (c) pressure, (d) GHSV.

Table 3.6. Distribution of the feed compositions.

Feed gas	Range (vol. %)
H_2	25.0 – 95.2
CO	0 – 66.7
Inert	0 – 50.0
CH_4	0 – 18.0
CO_2	0 – 45.1
H_2O	0 – 33.0

The type of reactor was also chosen as the last input variable in the database. It is a categorical variable comprising fixed-bed, slurry, and microchannel reactors. In order to reduce the dimensionality of the database, fluidized bed and continuous-stirred tank reactors were embedded in slurry type reactors because of their similar structures. Analogously, monolith reactors were considered as microchannel reactors. Figure 3.10 shows the reactors used in the experiments with respect to the number of data points.

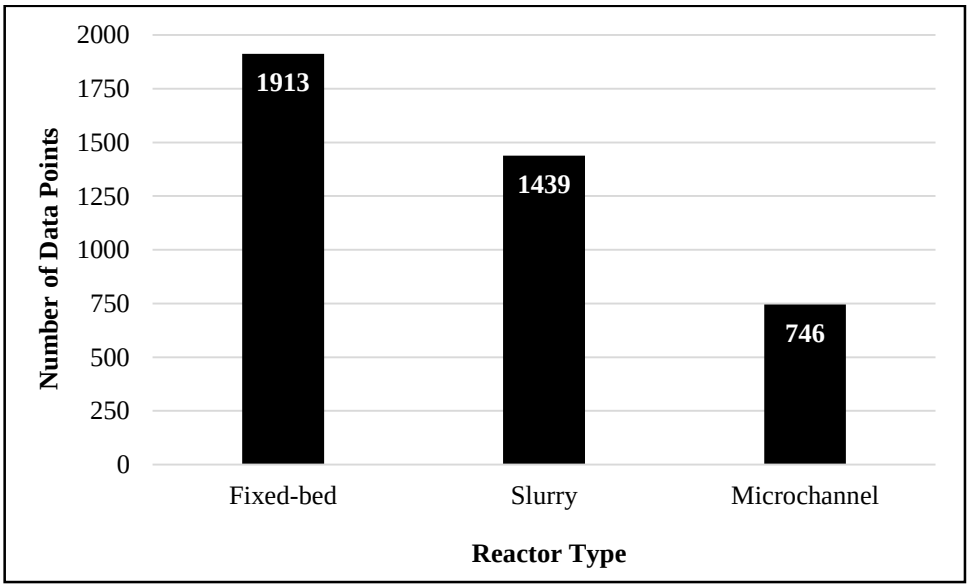

Figure 3.10. Distribution of the reactor types.

The output variable, CO conversion, was a continuous attribute, reported between 0 and 100 in per cent. Figure 3.11 shows the distribution of the CO conversion with respect to the number of data points.

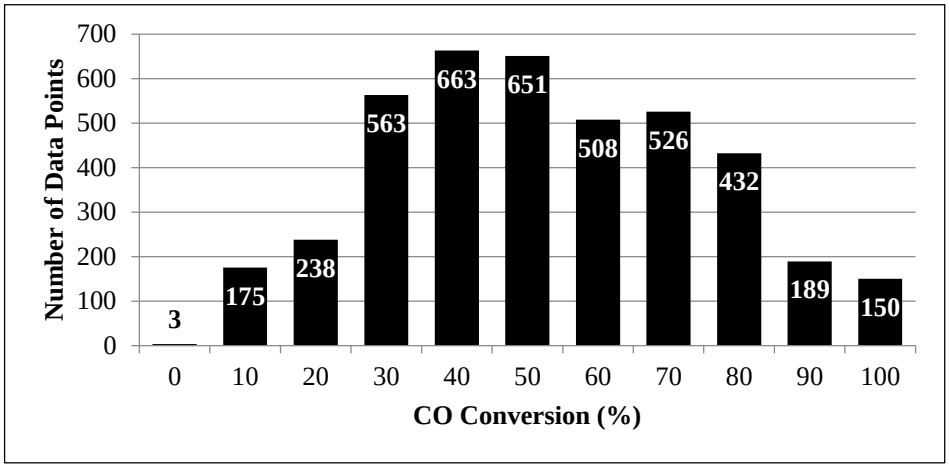

Figure 3.11. Distribution of the CO conversion.

3.2. Modelling the Reaction

The computational work was carried out in R 3.1.2 [223], by using the environment provided by RStudio [224]. Various supervised and unsupervised learning algorithms were used to extract knowledge in various kinds from the database, construct models to predict unstudied experiments, and hopefully guide future studies to develop improved catalysts and find optimum operating conditions. The computational work started with preprocessing the original database. Subsequently, PCA was employed for a descriptive approach to reveal hidden relationships between the input variables and the output variable, and also determine significant variables. Multiple linear regression, ANN, DT, and RF models were used to predict the numerical value of CO conversion.

3.2.1. Preprocessing

As mentioned in Section 3.1., the database contained missing values in specific attributes, and the range of attributes varied drastically depending on the content of the attribute. Preprocessing the original data was initiated by completing the missing values in catalyst preparation related variables. Calcination time and reduction time were two attributes that some of the researchers had missed to report in their articles. Although these two attributes were numerical and continuous, they were mostly integers since experimenters prefer calcining and reducing their catalysts in discrete values of time such as 3 h, or 4 h, instead of a value like 3.26 h. Therefore the most frequently used values, i.e. the modes of

the attributes, were assigned to complete the missing values. TOS was another parameter with occasional missing values. Similar to the case with calcination time and reduction time, researchers prefer reporting their results at steady state, which is usually a discrete value such as 12 h, or 24 h. Accordingly, the mode of TOS variable was assigned to the missing values.

Surface area, pore volume, pore diameter, and metal particle size were the last attributes with missing values. These physical properties were functions of the catalyst preparation methods, as previously mentioned. Therefore, instead of classical approaches like the statistical methods, machine learning algorithms were applied to predict these unknown properties from the catalyst synthesis variables. The number of known and unknown instances of each attribute was given in Table 3.7.

Table 3.7. Number of known and unknown instances.

Attribute	**Number of known instances**	**Number of unknown instances**	**Fraction of known instances (%)**
Surface Area	3534	564	86.2
Pore Volume	3079	1019	75.1
Pore Diameter	2738	1360	66.8
Metal Particle Size	1517	2581	37.0

ANN algorithm, provided by RSNNS package [225] was used to construct a mathematical model for each physical property. Catalyst preparation related variables, namely base metal loading, promoter loading, support type, base metal loading technique, promoter loading technique, calcination conditions, and reduction conditions were used as the input variables of the model. Additionally, the physical properties were determined to be weakly correlated with each other as mentioned above. Hence, while predicting one of these properties, other properties were also included in the input set. Prediction of each property was carried out one by one with separate models, and in the order of decreasing fraction of known instances which were provided in Table 3.7. Every separate model included the preceding property as an additional attribute. In other words, the model for surface area, for instance, included merely the catalyst preparation related variables, whereas the pore volume

model also included the surface area as an input variable as well; and pore diameter model contained both surface area and pore volume in the input variables. The reason for applying such an indirect route instead of directly constructing the models only with catalyst preparation related variables was to benefit from the mild correlations between the physical properties, while minimizing the accumulating uncertainty as the algorithm proceeds to farther properties. The complete algorithm followed during this process was visualized in Figure 3.12.

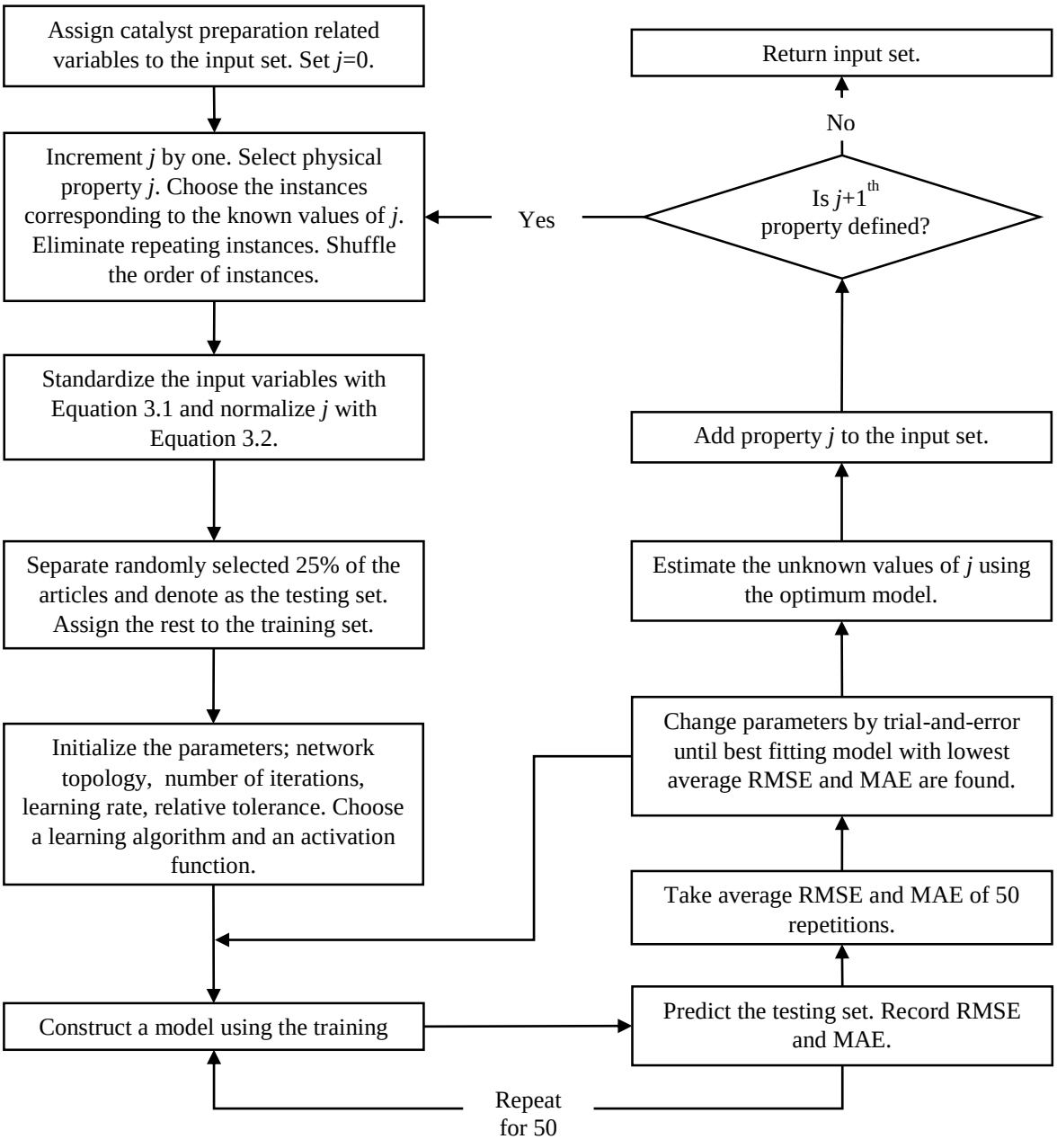

Figure 3.12. Process diagram of missing data completion in physical properties.

The process started with assigning previously mentioned catalyst preparation related variables as the input estimators. Since surface area had the greatest fraction of known instances, the algorithm started with constructing models for that property. The instances with known values of surface area were selected to construct the model for surface area prediction. Most of these instances had duplicates in the database since some catalysts were tested under different reaction conditions. These duplicates were eliminated to avoid dominance of these catalysts on the model. The estimator variables in the dataset were standardized by Equation 3.1 to prevent dominance of the attributes with wider range to those with narrower range.

$$z = \frac{x - \mu}{\sigma} \tag{3.1}$$

Where, z is the standard score, x is the value that the instance takes at a specific attribute, μ and σ are the mean and standard deviation of the attribute, respectively. The target variable, surface area, was mapped between 0 and 1 by normalization, described by Equation 3.2.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{3.2}$$

Where, x' is the transformed value, x is the value that the instance takes at a specific attribute, x_{min} and x_{max} are the minimum and maximum values of the attribute, respectively.

25 % of the research papers were randomly selected and separated from others to be tested after training the model. The rest of the instances were labeled as training set. One should note that, instead of splitting the dataset according to the number of instances, the number of research papers was taken as the splitting basis to avoid spoiling the data. Required parameters, which were the number of hidden layers and neurons, activation function, learning algorithm, maximum number of iterations, learning rate, and relative tolerance were initialized randomly to train the ANN model. The ANN model was constructed by using the training set and the initialized parameters. The validity of the model was tested by predicting the testing set. Root-mean-square error (RMSE) and mean absolute

error (MAE) were calculated to observe the accuracy of the model, as described in Equations 3.3 and 3.4, respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.3}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3.4}$$

Where, n is the number of instances in the testing set, y_i and $\hat{y}_i$ are the observed and predicted values of the ith instance, respectively.

Since ANN algorithm randomly initializes the weights and biases of the network, it yields different results when constructed several times. In order to be able to evaluate the models statistically, training and testing steps were repeated for 50 times. Average RMSE and MAE values were calculated for the repeated models.

Different models were constructed by changing the parameters and the network topology. The network giving the lowest RMSE and MAE was assumed to be the global minimum, and corresponding coefficient of determination was computed by Equation 3.5.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3.5}$$

Where, R^2 is the coefficient of determination, and $\bar{y}$ is the mean of the observed values. The selected parameters were used to train the complete input set. Unknown instances were predicted with the resultant model and hence, the missing values of the attribute were completed. The completed attribute was added to the input set and the next property was predicted with the same procedure until all four attributes were completed.

After completing all the missing variables in the entire database, the input variables were standardized by using Equation 3.1, and the output variable, i.e. CO conversion was mapped between 0 and 1 by Equation 3.2.

In addition to completing the missing variables and standardizing, the dataset was split into two subsets regarding the base metal. The instances containing Co in the catalyst were labeled as Co-based set and similarly, Fe containing instances were labeled as Fe-based set. Naturally, bimetallic catalysts were included in both of the models. A Ru-based subset was also created, but contained too few instances to construct a generalizable model. Hence, it was neglected while constructing the models. The complete dataset was also divided with respect to the distribution of the operating temperature. The instances with an operating temperature below 520 K were assigned to the low temperature (LT) set, and the rest to the high temperature (HT) set. The reason for selecting a rather arbitrary boundary instead of the true boundary between LTFT and HTFT was to balance the number of instances between two subsets.

3.2.2. Knowledge Extraction with Unsupervised Learning Algorithms

Before constructing the models with supervised machine learning algorithms, a preliminary analysis was carried out by principal component analysis (PCA) to acquire a prior knowledge on the database. "prcomp" function from the "stats" package was used for this purpose..

Two subsets, namely operating conditions (TOS, temperature, pressure, GHSV, H_2 vol.%, CO vol.%) and physical properties (surface area, pore diameter, pore volume, metal particle size), were created to compare with CO conversion. Each subset was subjected to PCA with CO conversion to observe any correlation. The first two principle components were plotted on a two dimensional biplot and the correlations were analyzed.

3.2.3. Modelling with Multiple Linear Regression

25 % of the instances were randomly separated and labeled as the testing set. The remaining instances were used to train a multiple linear regression model, as described in Section 2.2.1. "lm" function in the "stats" package [223] was used to train the model and determine the regression coefficients. The validity of the model was tested by introducing the testing set to the model.

3.2.4. Modelling with Artificial Neural Network

R environment provides several packages for ANN implementations. “nnet”, “AMORE”, “neuralnet”, “monmlp”, “qrnn”, “frbf”, and “RSNNS” are some of the successful ANN packages. In this work, RSNNS package [225] was preferred since it provides an easily comprehensible high-level interface, while enabling modification on the network topology, the learning algorithms, and activation functions. Additionally, “NeuralNetTools” package [226] was used to determine the important variables of the model and conduct a sensitivity analysis.

The process for constructing a neural network was summarized and visualized in Figure 3.13. It started with labeling the instances with the same experimental run. This step was mandatory, and the aim was to prevent spoiling the testing and training sets while splitting. In other words, even if one instance from an arbitrary experimental run had appeared in the training set, it would have given a crucial hint to predict other instances from the same experimental run. Since the aim of the predictive approach is to estimate unstudied experiments, spoiling the testing and training sets would violate the aim of this work.

After labeling the instances, 25 % of the experimental runs were randomly selected as the testing set, and the remaining was assigned as the training set. Subsequently, the training set was also divided into 10 folds with respect to experimental runs, for a 10-fold cross validation. Cross validation was used as a validation method to evaluate the complete training set with the same parameters. Detailed procedure of cross validation was explained along with the network construction process.

Before training the model, a network topology and size were assumed with the number of hidden layers and neurons. Along with the network size, learning algorithm and activation function was decided. The initial guesses were assigned for relative tolerance, learning rate and momentum

One of the 10 folds was separated from the training set, and remaining folds were used to train a network by using “mlp” function of RSNNS package. Relative importance of each attribute in the model was determined by using “garson” function of the NeuralNetTools package, which utilizes an algorithm proposed by Garson [227]. 5 of the least important

attributes were recorded. Previously separated fold was introduced to the model to test the accuracy of the model. This training-testing cycle was repeated for each fold. RMSE and MAE values were calculated for the complete training set by Equations 3.3 and 3.4.

The attributes which were consistently recorded as one of the 5 least important attributes were removed from the dataset. Removal of these attributes significantly helped reducing the dimensionality problem. Accordingly, the model was once again constructed by 10-fold cross-validation. This cycle of constructing model and removal of the redundant attributes was pursued as RMSE and MAE kept decreasing.

After finding the optimal set of attributes; the learning rate, relative tolerance and momentum were also optimized by trial-and-error. The combination of these parameters which yielded the minimum RMSE and MAE was selected as the best parameters. These parameters were used to train a model by using the complete training test. Testing set was introduced to the resultant model. RMSE, MAE, and R^2 values were calculated to measure the accuracy of the model. Standardized residuals of the testing set were computed by Equation 3.6 to conduct a residual analysis.

$$z_i = \frac{\varepsilon_i - \bar{\varepsilon}}{\sigma} \tag{3.6}$$

Where z_i is the standardized residual, ε_i is the difference between the observed output and predicted output, i.e. the residual, and $\bar{\varepsilon}$ and σ are the mean and standard deviation of the residuals.

The importance analysis was carried out with “garson” function, and a sensitivity analysis was conducted on 3 of the most important attributes by “lekprofile” function of the “NeuralNetTools” package, which implements a sensitivity analysis by using Lek’s profile method [228]. The sensitivity analysis was used to analyze the changes in CO conversion when a unit change is applied on a specific input attribute, while keeping other variables at maximum, minimum or in-between levels.

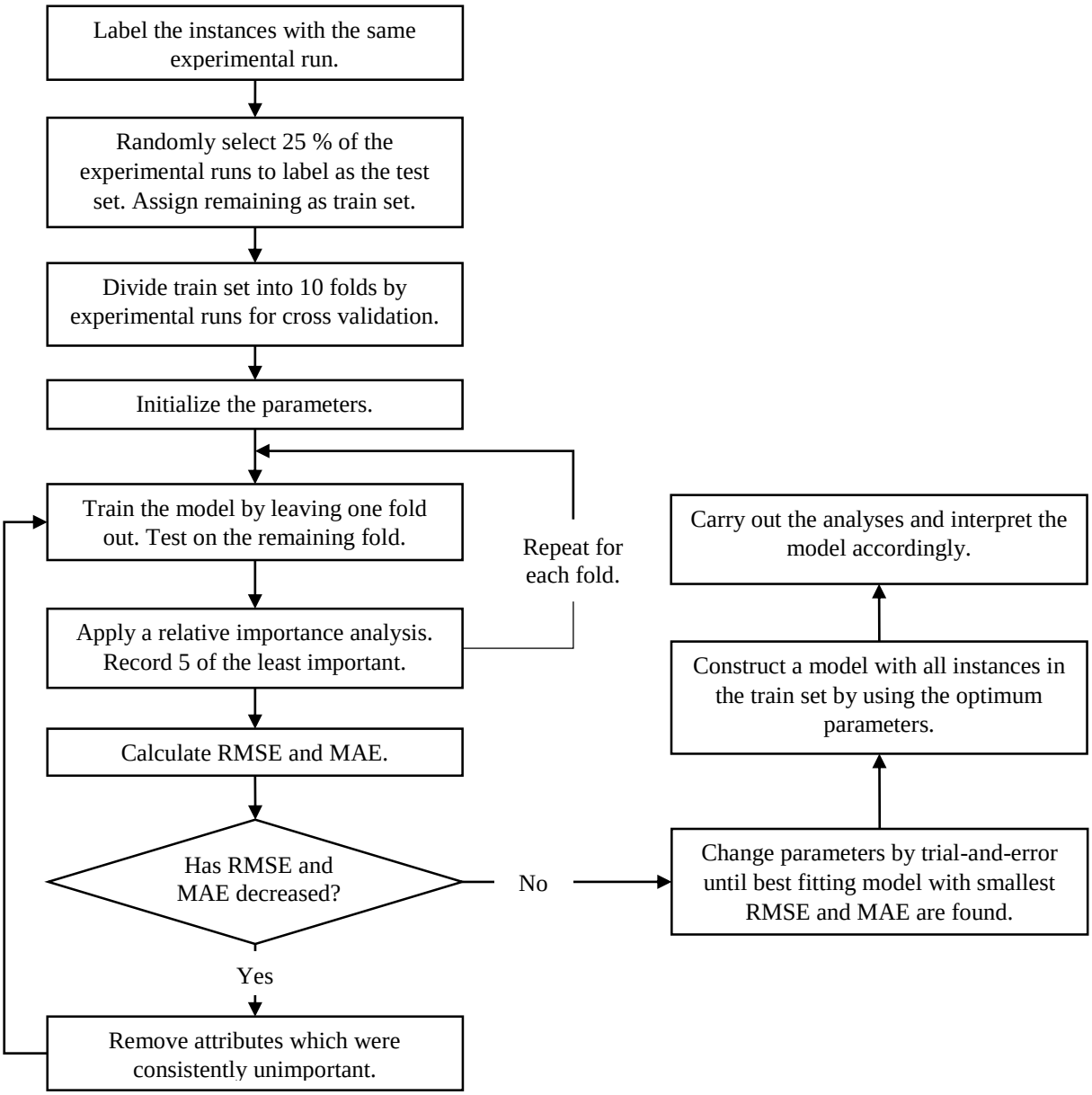

Figure 3.13. Neural network construction process.

3.2.5. Modelling with Decision Tree

There are a few packages for classification and regression tree applications, such as "rpart", "party", and "tree". Although all decision tree packages yielded similar results, "rpart" package [229] was used in this work since it was the most widely used package.

Decision tree algorithm was used for classifying discrete FTS performances. Accordingly, the output variable, CO conversion, was discretized into four classes; 0 – 30 (very bad), 30 – 50 (bad), 50 – 75 (good), 75 – 100 (very good). The class intervals were

intentionally selected asymmetrically, in order to balance the number of instances in each class as much as possible while keeping them meaningful and interpretable.

After classifying the instances, 25% of the instances were randomly selected to be tested after training the tree. The remaining 75% was split into 10 folds for cross-validation. “rpart” function was used to grow the trees, and “rpart.control” function was used to tune the decision tree parameters. The “minbucket” argument was adjusted to control the number of observations in the terminal nodes. “cp” was the complexity parameter, indicating the minimum allowed decrease in overall lack of fit after splitting. “maxcompete” was the maximum number of competing splits retained in the output. Lastly “maxdepth” was the maximum depth of any node in the resultant tree. The last argument was especially kept as low as possible to construct an interpretable tree. Yet, since the algorithm failed to produce a fruitful, meaningful tree with high accuracy, the algorithm was used as another black box machine learning method. In other words, “maxdepth” argument was increased to higher values, which increased the predicting power of the trees while diminishing the interpretability.

3.2.6. Modelling with Random Forest

“randomForest” package [230] was used to model the FTS performance dataset with random forest algorithm. The process of constructing and validating random forest models was very similar to ANN models and hence, the same process diagram in Figure 3.13 was followed.

“randomForest” function was used to grow the forest. Analogous to the ANN model, the forest size, the argument “ntree” in the function, was initialized prior to growing the forest. In addition, number of variables randomly sampled as candidates at each split, the argument “mtry”, was assigned an initial guess. After the forest was grown, the importance analysis was carried out by using “importance” function. This function returns two measures of importance for regression problems, i.e. mean decrease in accuracy and mean decrease in node impurity. In this work, the former was employed as a measure of importance. This value was computed from permuting out-of-bag data. Mean-square-error was calculated from prediction of the out-of-bag data. The same procedure was repeated after permuting

each predictor variable. The difference between the two are averaged for all trees and then standardized.

The rest of the procedure was the same with ANN. The best representing variables were chosen, parameters were optimized and used to train the complete training set. RMSE and MAE were calculated for the testing set to measure the accuracy, residual and importance analyses were carried out.

4. RESULTS AND DISCUSSION

FTS performance was empirically modelled based on the preprocessed database by utilizing descriptive and predictive approaches. PCA was employed for an unsupervised descriptive approach. On the other hand, predictive models were constructed by using multiple linear regression, ANN and random forest to estimate the numerical value of CO conversion. Additionally, decision tree algorithm was applied to address this regression problem as a classification problem in order to deduce meaningful, physically interpretable and applicable rules.

4.1. Completing Missing Physical Properties

Completing the missing values in surface area, pore diameter, pore volume, and metal particle size was a complementary task prior to modelling the CO conversion. The procedure described in Figure 3.12 was followed to predict the missing values of physical properties.

Surface area was the first physical property to be predicted, since it contained the least number of unknown instances, as shown in Table 3.7. 25% of the articles with known surface area were randomly selected and tested by the model trained by the remaining 75%. 541 unique instances remained in the train set after eliminating the duplicates, and 154 instances were used to test the accuracy of the resultant models.

While constructing the models, one to 12 neurons in a single hidden layer were tried to find the best fitting network size. Additionally, the investigated learning span was from 0.01 to 0.90, absolute tolerance span from was 10^{-9} to 10^{-1}, and maximum number of iteration span was from 10 to 3000. After repeating each candidate model for 50 times, the optimum network was obtained with 8 neurons, while the learning rate was 0.6, absolute tolerance was 10^{-6}, and maximum number of iterations was 1500. The activation for all hidden units was performed by “Act_IdentityPlusBias”, which activates the corresponding unit with the weighted sum of all incoming connections, and adds the bias of the unit. The accuracy of the best fitting model with respect to testing data was demonstrated in Figure 4.1.

It was observed that the model can accurately predict the surface areas between 0 and 300 m^2/g with good approximations; the values between 300 and 600 m^2/g were also predicted with tolerable differences. Nevertheless, surface areas larger than 900 m^2/g were estimated to be around 400 – 500 m^2/g. The reason for this deviation from satisfactory predictions at high surface areas might stem from a class imbalance problem. The number of instances with high surface areas was probably insufficient in the training set, which caused poor learning in those regions. This claim might be reinforced with the distribution of surface area in the dataset, as shown in Figure 3.8. It can be seen that the number of instances with a surface area higher than 500 m^2/g comprised only about 2.5% of the complete data.

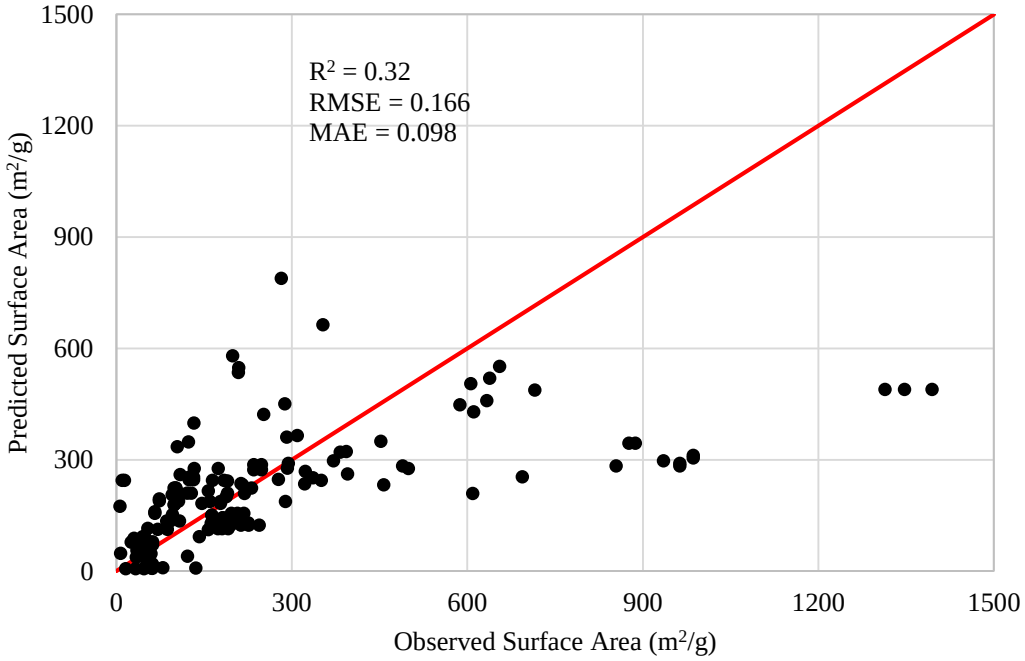

Figure 4.1. Predicted vs. observed surface area with testing data.

Still, this model was found to be satisfactory because of its predictive power at lower surface areas. Hence, the optimum parameters were used to train the other models with the complete training set to estimate other physical attributes and ultimately the CO conversion.

The subsequent step was completing the missing values of pore volume. After splitting 25% of the articles, 436 instances remained in the training set, and the testing set consisted

of 130 instances. The previously stated parameter intervals were investigated to search for the optimum settings. The best fitting model was maintained with 2 neurons in a single hidden layer. The learning rate, absolute tolerance, and maximum number of iterations were 0.1, 10^{-5}, and 75, respectively. Figure 4.2 shows the accuracy of the resultant model on the testing set.

Similar to surface area prediction model, pore volume model also estimated the instances with lower pore volumes with higher accuracy. It was observed that the predicted values are close to the observed values for pore volumes smaller than 0.4 cm^3/g. This result was plausible since the majority of the instances lie within that region and enhance the learning ability by creating a bias. However, the rest of the instances were quite problematic due to poor sampling, and hence they were predicted with lower accuracy. Yet, the model appeared to be available for prediction of the missing data since the order of magnitude of the predictions were approximate to observed values.

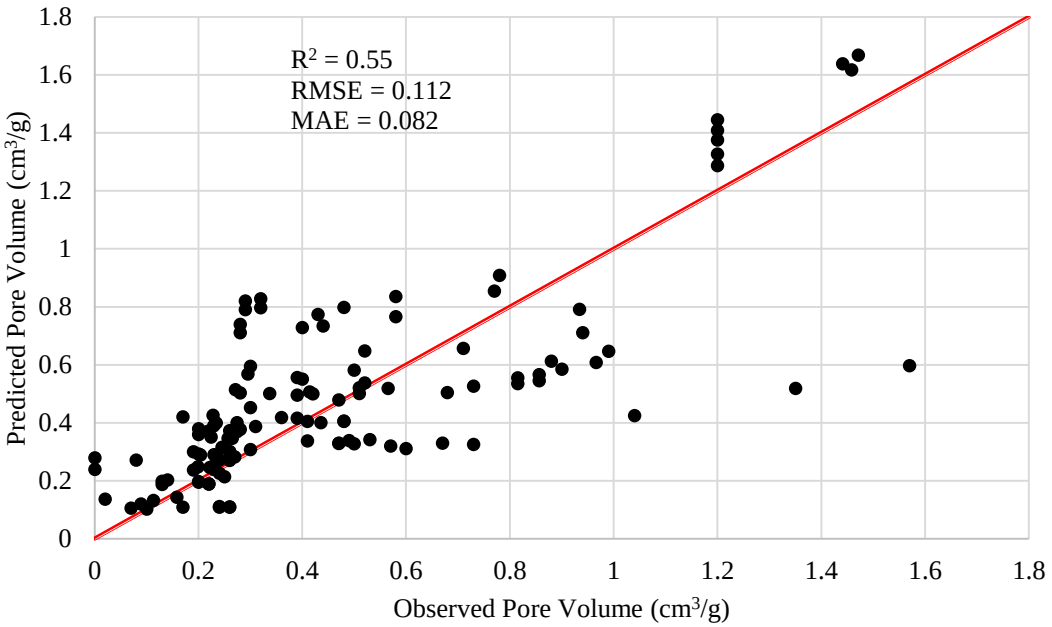

Figure 4.2. Predicted vs. observed pore volume with testing data.

The same procedure was conducted for the remaining physical attributes, pore diameter and metal particle size. The number of instances in training and testing sets along with the optimum parameters for the two models were given in Table 4.1.

Table 4.1. Optimum parameters for pore diameter and metal particle size models.

	Pore Diameter Model	**Metal Particle Size Model**
Number of instances in the training set	594	458
Number of instances in the testing set	177	89
Number of neurons	3	2
Learning rate	0.6	0.8
Absolute tolerance	10^{-7}	10^{-7}
Maximum number of iterations	50	100

It was observed in Figure 4.3 that pore diameter model can successfully predict the testing data both in lower and higher regions, although the number of observations was limited in the former. The reason for this success was probably including other physical attributes in the dataset after completion. Since surface area and pore volume have a strong physical association with pore diameter size, they might have improved the model significantly.

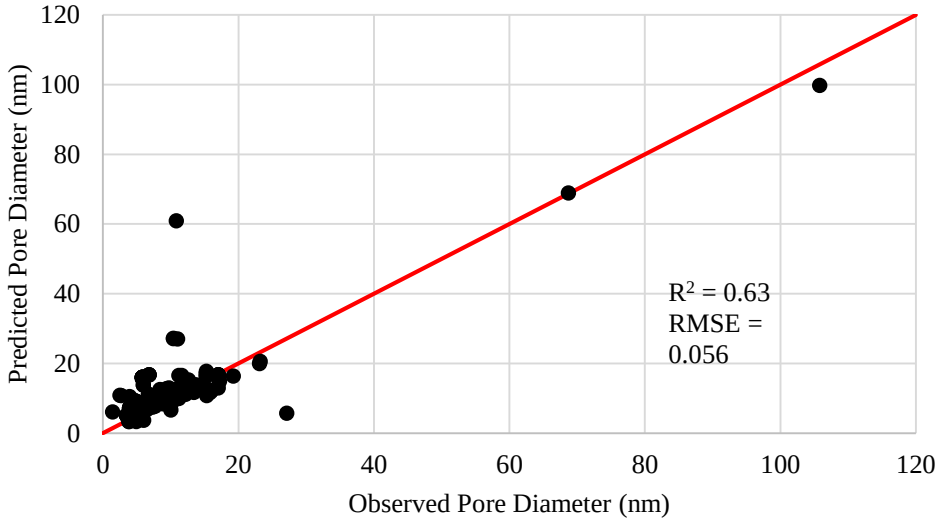

Figure 4.3. Predicted vs. observed pore diameter with testing data.

Figure 4.4 shows that prediction of metal particle size was much more successful than of the preceding physical properties. The predicted values mostly tended to lie on the identity line except a few outliers. Also, a bias towards higher values was observed at metal particle sizes below 5 nm.

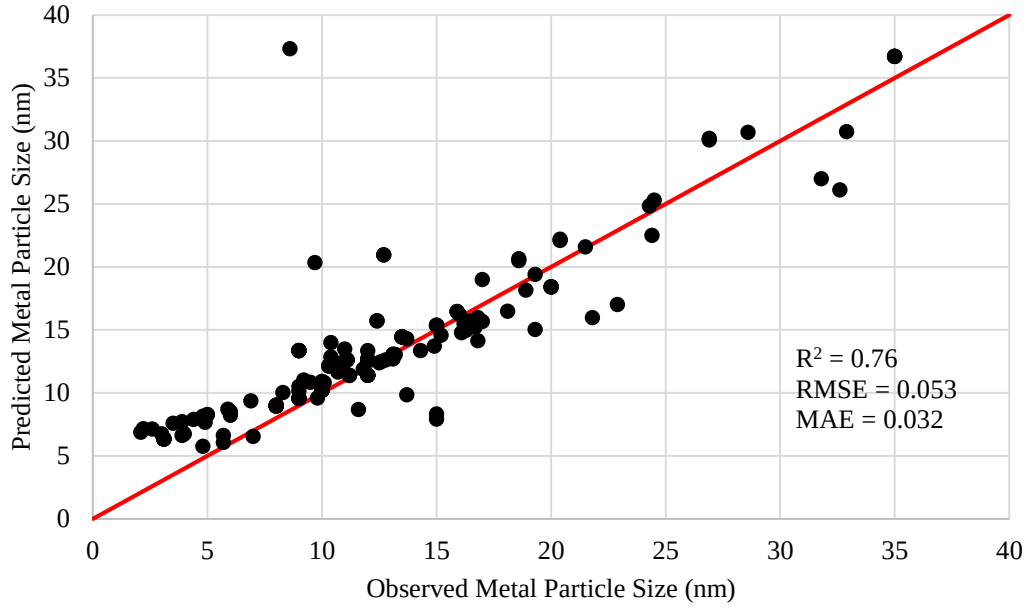

Figure 4.4. Predicted vs. observed metal particle size with testing data.

Consequently, all the missing values were completed by employing the last model on missing metal particle size values. The models used in the preprocessing step were constrained by the input variables and instances in the dataset and thus, they would fail to capture any unique and novel catalyst preparation techniques. Yet, they provided satisfactory approximations to their corresponding testing sets, indicating that physical properties can be successfully modeled by ANN as an additional task.

4.2. Knowledge Extraction by Principal Component Analysis

PCA was used to reduce the dimensionality of the preprocessed data to map the input and output variables on a two dimensional plot. Yet, instead of introducing the complete dataset at once; the two subsets, comprising the operating conditions and the physical properties as mentioned in Section 3.2.2, were compared with CO conversion to obtain more interpretable images.

Figure 4.5a shows the correlation between the physical properties and CO conversion on the first two principal components. The small angle between CO conversion (CO.Conv) and pore diameter (P.D.) indicated a strong relationship, whereas the right angle between CO conversion and surface area (S.A.) and pore volume (P.V.) implied negligible relationship. On the other hand, CO conversion and metal particle size (M.P.S.) appeared to be mildly negatively correlated, indicating that small metal particles leads to higher FTS activity. This result was reasonable since smaller particles on the catalyst surface imply high dispersion of the active metal.

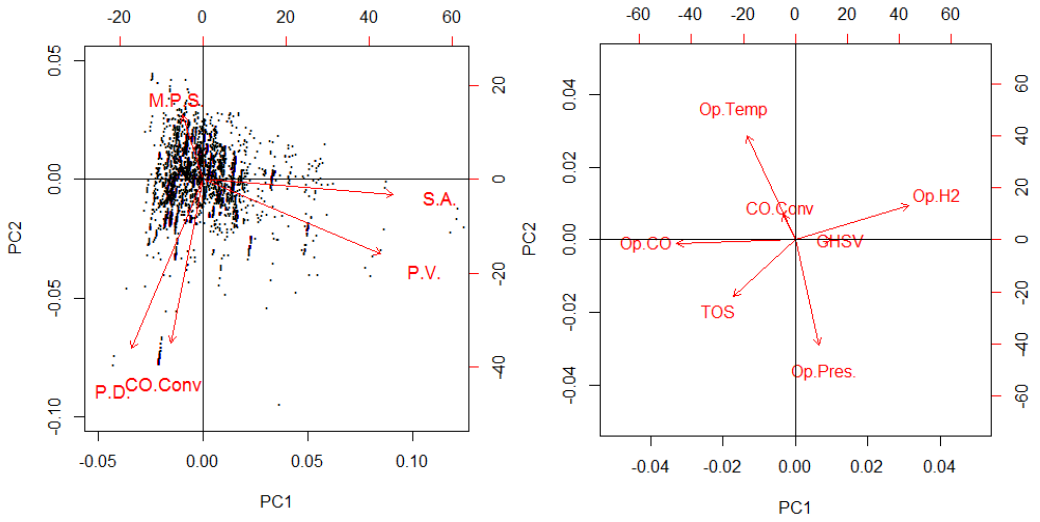

Figure 4.5. PCA biplot with (a) physical properties (b) operating conditions.

Figure 4.5b shows a similar analysis between the operating conditions and CO conversion. The data points in this figure were omitted to obtain an interpretable image, since they obscured the vectors. Operating temperature (Op.Temp) and CO conversion pointing in the same direction was construed as a very strong correlation. This was a plausible conclusion since increased temperature increases reaction kinetics in accordance with Arrhenius Equation until the thermodynamic limit. Additionally, this finding was reinforced by myriad studies in the literature, explicated in Section 2.1.5. The effect of hydrogen (Op.H2) and CO (Op.CO) fraction in the feed stream appeared to be negligible. GHSV was negatively correlated with CO conversion, which was an expected result. GHSV is an indicator of the amount of reactant treated in the unit reactor volume per unit time. Thus, it is expected for higher GHSV values leading to lower CO conversion values. Time on stream (TOS) was also found to have a mildly adverse effect on CO conversion. This

result can be explained by the decreasing stability of the catalysts with time. The FTS catalysts usually deactivate and lose their activity due to coke formation and catalytic poisoning with time.

Table 4.2. Variance decompositions of the PCA biplots.

	Physical Properties		**Operating Conditions**	
	PC1	**PC2**	**PC1**	**PC2**
Standard deviation	1.28	1.06	1.23	1.1
Proportion of variance	0.33	0.23	0.30	0.24
Cumulative Proportion	0.33	0.56	0.30	0.54

Table 4.2 shows the variances explained by the first two principal components for the two PCA biplots. It was observed that neither biplots were sufficiently significant since cumulatively explained variances were 56% and 54%, respectively. Still, they provided rather broad idea on the relation between physical properties, operating variables and CO conversion.

4.3. Prediction by Multiple Linear Regression Model

Multiple linear regression technique was the first method used as a supervised learning algorithm. The reason for preferring this technique was that it is easy to comprehend, intuitive, and widely used in different areas.

75% of the instances were separated as the training set and remaining 25% was used to test the accuracy of the model. Figure 4.6 shows the accuracy of the model on testing data. It was observed that the predictions were very far from being realistic. Therefore no further effort was spent to improve the model or to interpret the results.

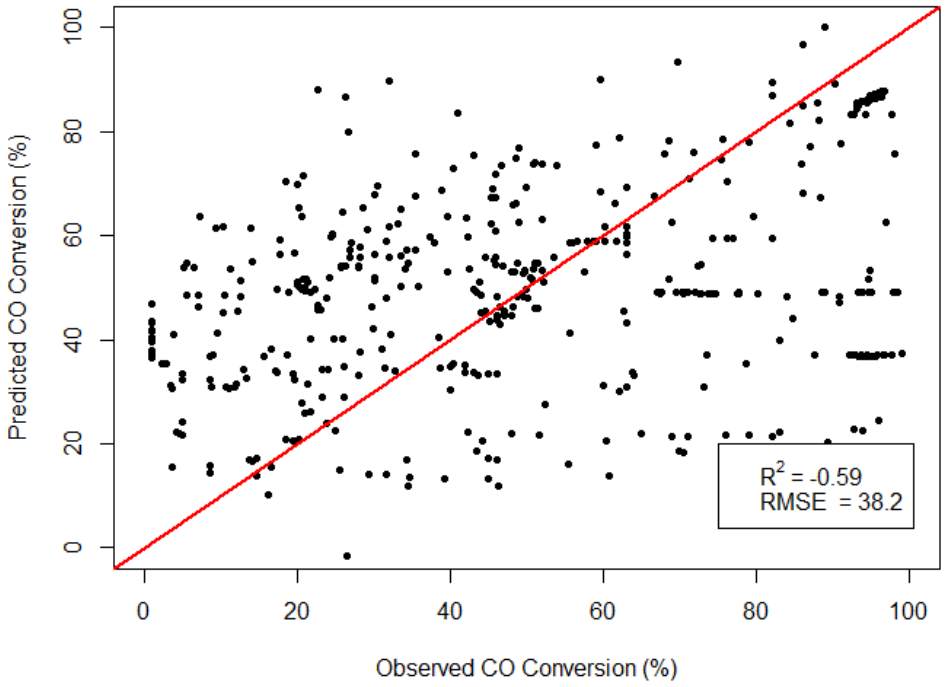

Figure 4.6. Predicted CO conversion vs. observed CO conversion by multiple linear regression.

4.4. ANN Prediction for Entire Dataset

The best fitting ANN model was determined by following the procedure described in Figure 3.13. RMSE was employed to compare the performances of the models, and the model with the lowest RMSE was chosen to be tested by the previously separated dataset. A wide range of number of neurons was tried one by one, starting from 5 to 120 neurons. 15 hidden units activated by the logistic sigmoid function (Act_Logistic in RSNNS package) and trained by standard backpropagation (Std_Backpropagation in RSNNS package) yielded the lowest RMSE and highest R^2 value. Although networks with a hidden layer size between 16 and 20 performed equally well, the resultant model was constructed by 15 hidden units, since it was less complicated. After maintaining the topology of the network, the parameters were optimized to acquire improved results. The learning rate, absolute tolerance, and maximum number of iterations were determined as 0.03, 10^{-9}, and 150, respectively.

Figure 4.7 shows the 10-fold cross-validation predictions of the previously separated training set with respect to their actual values. Each experimental run was marked with a different color as much as possible to observe whether the model captured the change in CO

conversion with respect to time on stream due to the characteristics of this attribute as mentioned in Section 3.2.4. It was observed in the figure that the model poorly estimates the instances in the validation folds. Especially most of the experimental runs marked with green were estimated to be constant during the course of reaction. Hence, it was deduced that the model was unable to capture the effect of TOS accurately. Apart from this problem, a bias was observed towards the range between 40 – 60% CO conversion. In other words, the model tended to estimate the instances in that interval. This bias might be explained by the majority of the observations piled over that region, as shown in Figure 3.11.

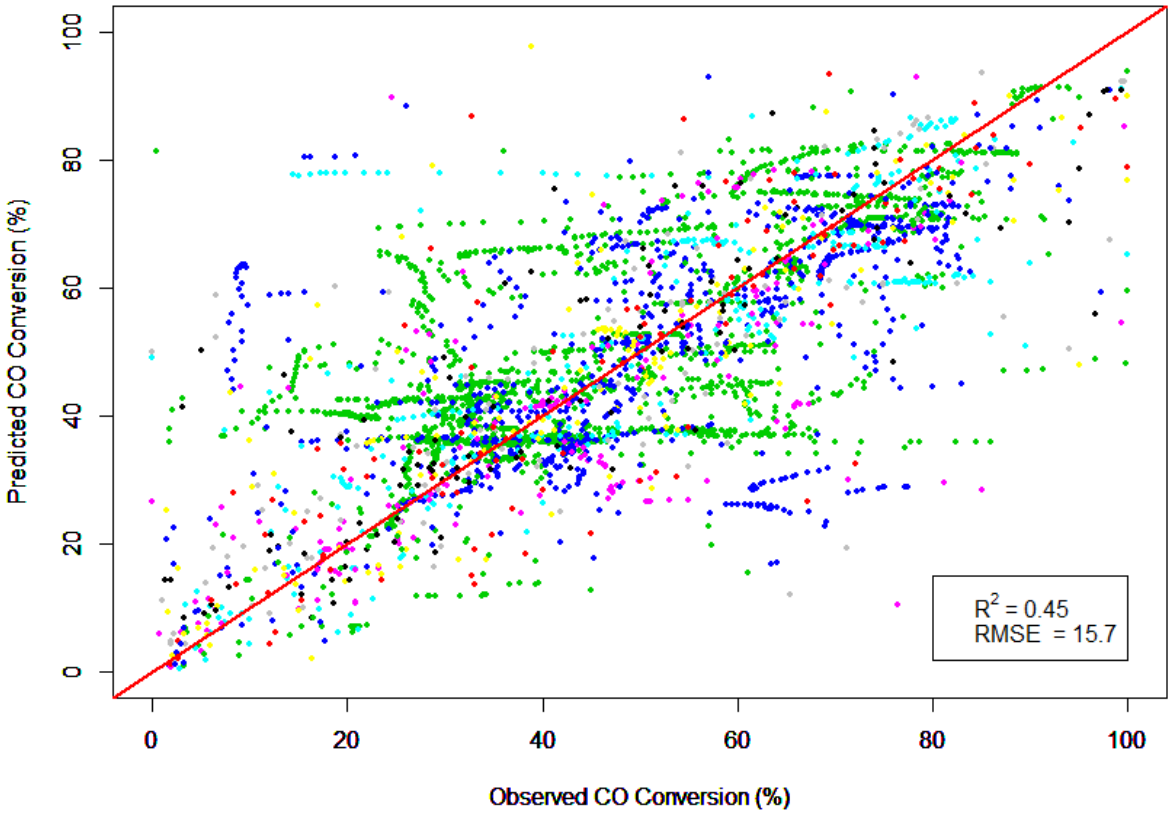

Figure 4.7. Predicted CO conversion vs. observed CO conversion for complete training set.

After the above mentioned parameters were heuristically optimized by cross-validating the training set, the model was validated by introducing the testing set. Figure 4.8 shows the accuracy of the model trained by the complete training set and optimum parameters. It was observed that the accuracy of the complete model was lower than the cross-validated models, regarding the R^2 and RMSE values. The most probable reason for the decrease in accuracy was that the parameters were tuned to minimize the error with

respect to the training set. Accordingly, the accuracy had decreased when an unseen set was introduced to the model.

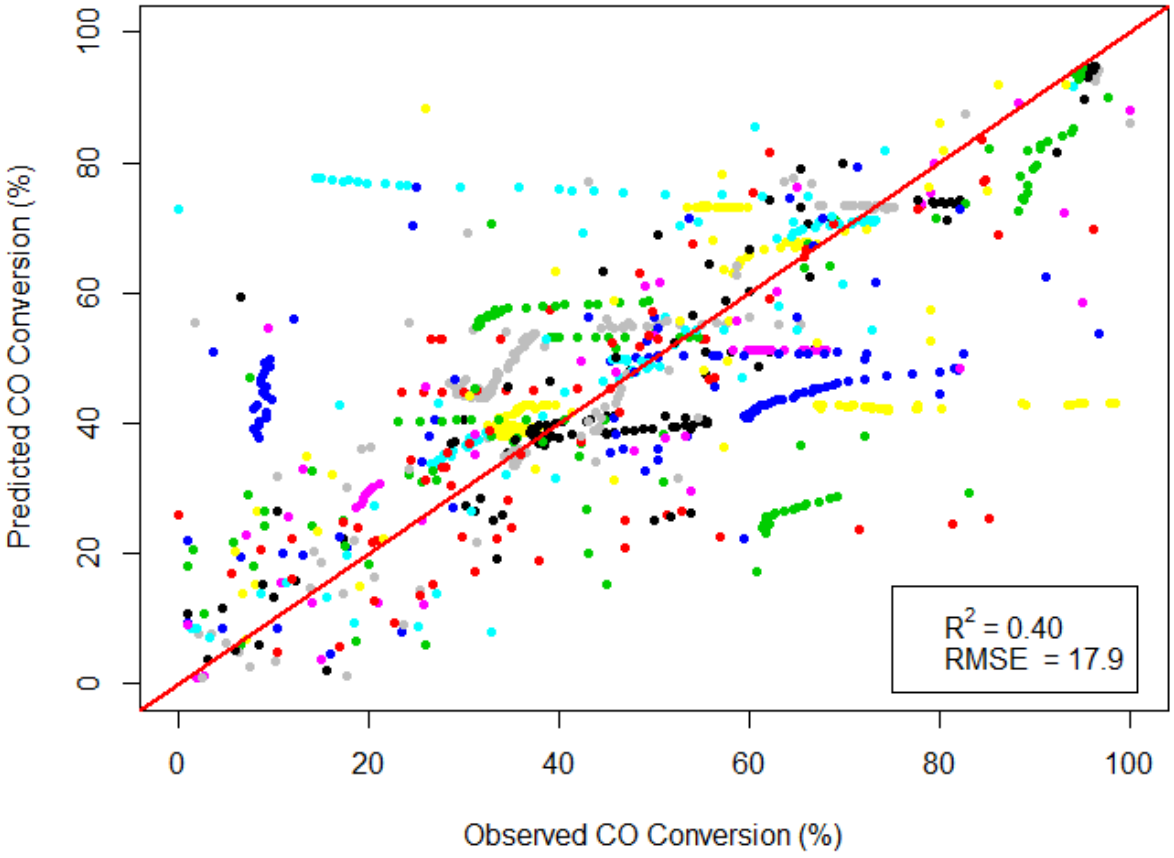

Figure 4.8. Predicted CO conversion vs. observed CO conversion for complete testing set.

The horizontal lines with the same colors indicated that the TOS dependent experimental runs were weakly modelled, since they were predicted to be constant throughout the experiment. Still, a few of the TOS dependent experimental runs were estimated with relatively greater success. Consequently, predictive power of the resultant model was far from being accurate, but sufficient to have a broad idea about the outcomes of certain catalyst design and operational variables.

4.4.1. Importance Analysis

The attributes that contributed constructing the model the most were analyzed by Garson's algorithm as shown in Figure 4.9. The relative importance of the attributes were plotted on polar coordinate system with decreasing order in clockwise direction. However,

only the first 36 attributes out of 52 were shown in the figure to keep it more simple and comprehensible. It was observed that operating temperature (Op.Temp) is the most important attribute contributing the model. This result was expected due to highly exothermic nature of FTS reactions. Along with operating temperature, pore volume (P.V.) and surface area (S.A.) ranked quite high among other attributes. Additionally, the composition of the reaction medium was determined to be an essential factor, since hydrogen (Op.H2) and inert fraction (Op.Inert) were important constituents of the model.

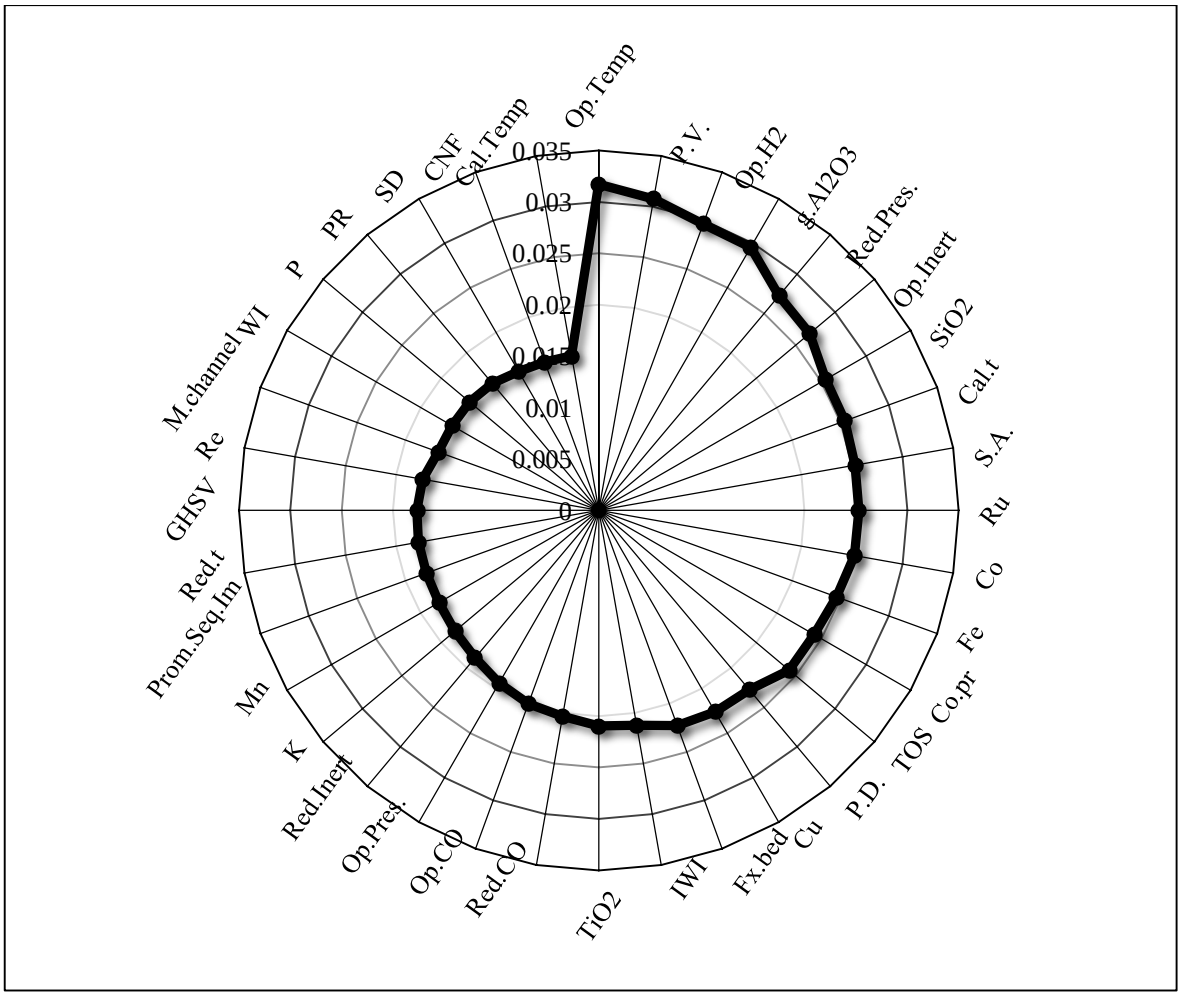

Figure 4.9. Relative importance of the attributes according to the complete training set.

It should be noted that the meaning of "relative importance" for the categorical variables in the radar chart is not as clear as that of continuous variables. For example, for computational easiness, all catalyst preparation methods were entered the database as separate variables (instead of the alternative levels for a single variables as "catalyst

preparation"), and the one used for each data point is labeled as "one" while the others had the value of "zero". If one catalyst preparation method is found to be important while the others are not, the result should be interpreted as follows: whether that method is used or not is important for the results but if that method is not used, it does not matter which other method is utilized. However, if these method have varying level of importance, the interpretation of the results may be more complicated. For the continuous variables, on the other hand, the relative importance simply shows whether the use of that variable in the model make a significant chances in results or not.

4.4.2. Sensitivity Analysis

In order to further analyze the effect of the important attributes, a sensibility analysis was carried out by Lek's profile method. The profile in Figure 4.10 shows the response in CO conversion across the range of three of the most important attributes, which were operating temperature, pore volume, and hydrogen fraction in the feed composition. All other input variables were held constant at the split fractions given on the right side. The split values were chosen to be the minimum, maximum, and at every 20^{th} percentile. Both the response and input variables were given in normalized form for convenience.

It was observed in Figure 4.10a that unless other input variables are at their extremes, an increase in the pore volume of the support increases the CO conversion. This result might be explained by mass transfer resistances in the narrow pores. Since the products of FTS are usually long hydrocarbon chains, larger pore volumes might enhance removal of the products from the active sites, increasing catalytic activity. Figure 4.10b shows the effect of operating temperature on CO conversion. It was observed that CO conversion tends to increase with increasing temperature, which was reinforced by the findings of PCA analysis in Figure 4.5b. Additionally, sensitivity analysis revealed that the positive effect of increasing temperature ceases after a certain point at low and middle split values. Lastly in Figure 4.10c, hydrogen composition in the feed stream was determined to be mostly ineffective. The only exception was that it significantly increased the CO conversion at 0.6 split value.

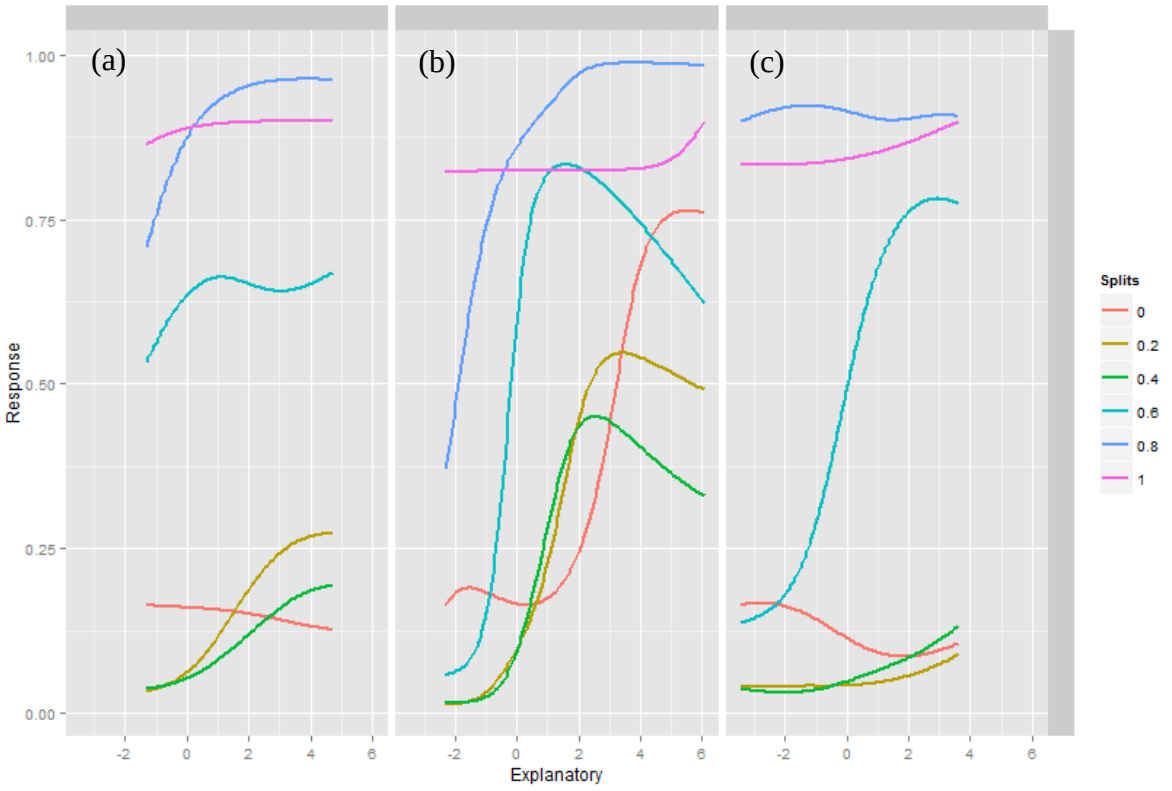

Figure 4.10. Sensitivity analysis for (a) pore volume, (b) operating temperature, (c) hydrogen composition.

4.4.3. Residual Analysis

A residual analysis was carried out on the testing results to reveal any patterns in the residuals with respect to the important attributes. A further motivation was to detect any outliers stemming from incorrectly modelling an attribute. Each experimental run was marked with a different color as much as possible to observe whether an experiment was consistently mispredicted due to a scaling problem.

Figure 4.11 shows the residuals of the resultant model with respect to operating temperature (a), pore volume (b), and hydrogen composition in the feed stream (c). The residuals were not plotted throughout the complete span of the attributes, but rather focused on the most densely populated regions. The residuals outside 99% confidence interval were defined as the outliers, hence +3 and -3 values in the residual plots were marked to identify the outlying predictions.

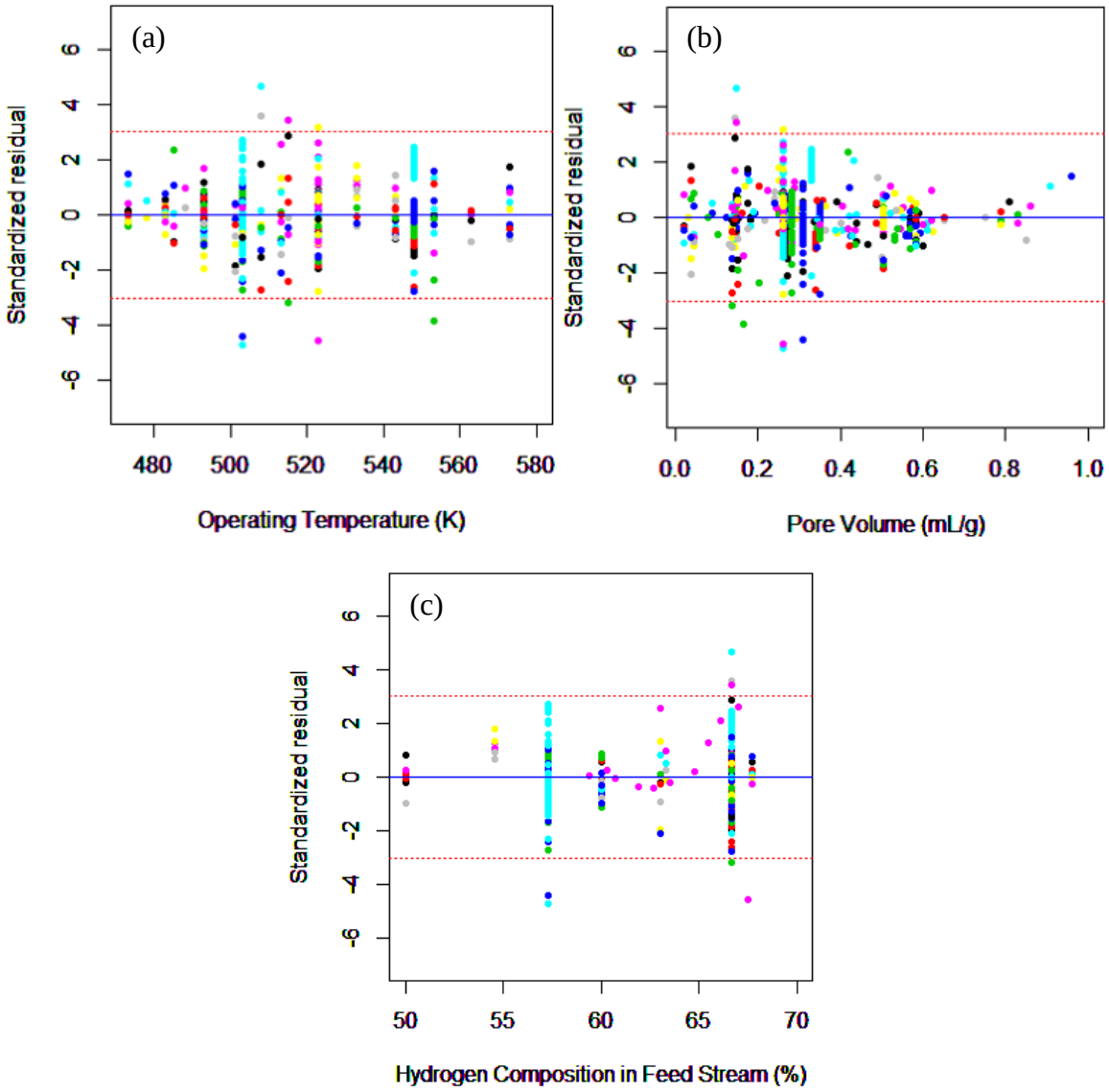

Figure 4.11. Standardized residuals vs. (a) operating temperature (b) pore volume (c) hydrogen composition in the feed stream.

It was observed that the residuals were randomly distributed around zero and no patterns were detected neither regarding the complete testing set, nor the individual experimental runs. This finding indicated that the model successfully captured the behavior of these significant attributes.

The residuals were also plotted against time on stream in addition to the most important attributes, as shown in Figure 4.12. It was observed that although the residuals were randomly distributed regarding the complete testing set, a few patterns were encountered

when evaluated by the experimental runs. The patterns colored with cyan, green and red were consistently lied on one side of the mean value. This result validated our a priori assumption, which posited that the time on stream variable was modeled poorly.

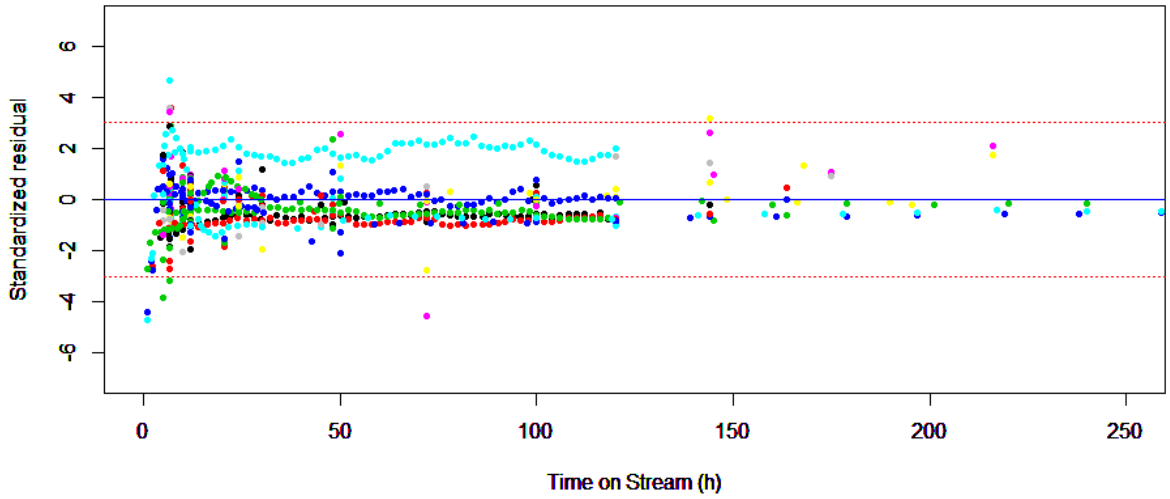

Figure 4.12. Standardized residuals vs. time on stream.

4.5. ANN Prediction for Individual Subsets

Apart from dealing with the complete dataset, different models were also constructed with the subsets involving Co-based, Fe-based, LT, and HT, which were mentioned in detail in Section 3.2.1.

4.5.1. Co-based Subset

The best fitting ANN model for Co-based subset was determined by the procedure described in Figure 3.13 with the intervals given in Section 4.3 for the parameters. The subset comprised 2069 instances in a 90 dimensional subspace. The optimum network was obtained with 21 units and a bias in a single hidden layer. Logistic sigmoid function was used for activation and standard backpropagation algorithm was used for learning. The learning parameters were tuned by cross-validating the instances and hence, the learning rate, absolute tolerance, and maximum number of iterations were found as 0.05, 10^{-8}, and 300, respectively.

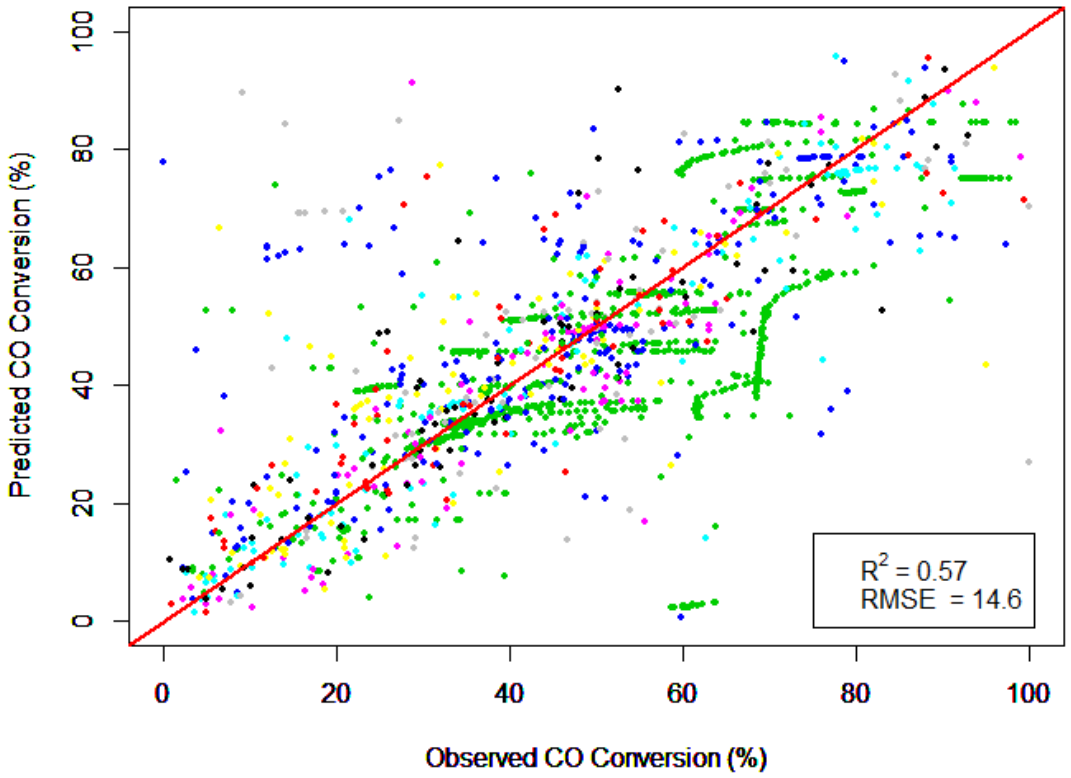

Figure 4.13. Predicted CO conversion vs. observed CO conversion for Co-based training set.

Figure 4.13 shows the predictions obtained by 10-fold cross-validation with the optimum parameters. It was observed that developing model specific to Co-based catalyst significantly improved the predictive power of the network regarding the increased R^2 value. Still, some of the TOS dependent experimental runs were observed to be fixed to a constant during the entire range of TOS. On the other hand, the bias observed in Figure 4.7 appeared to be less dominant in this model, which contributed to the higher predicting power of the model.

The accuracy of the model on the testing set was demonstrated in Figure 4.14. The performance of the model had improved when compared with the model trained with the complete dataset. R^2 value had increased from 0.40 to 0.50 and RMSE had decreased from 17.9 to 16.6. It was observed that TOS dependent instances were problematic in this model as well, but the general trend was captured in a few of them. In conclusion, Co-based catalyst performance was estimated with higher accuracy with this model.

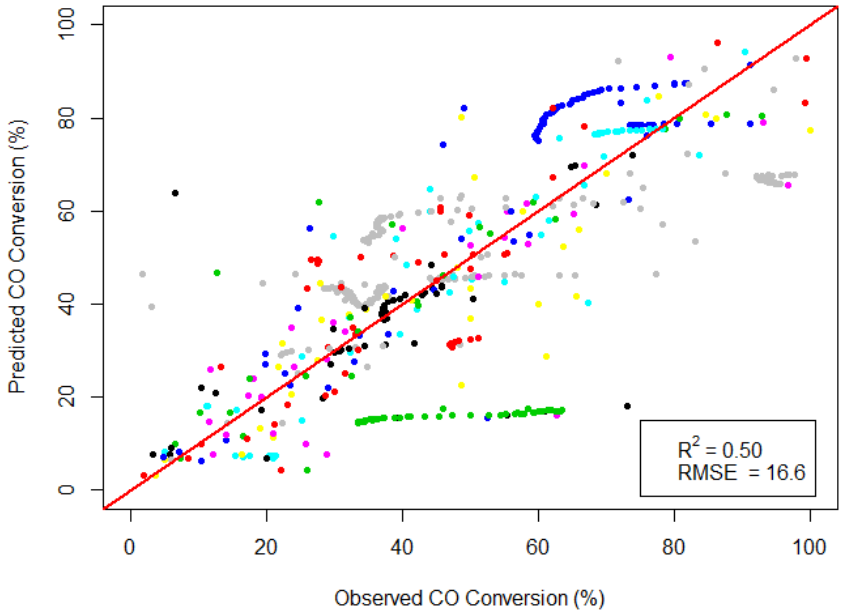

Figure 4.14. Predicted CO conversion vs. observed CO conversion for Co-based testing set.

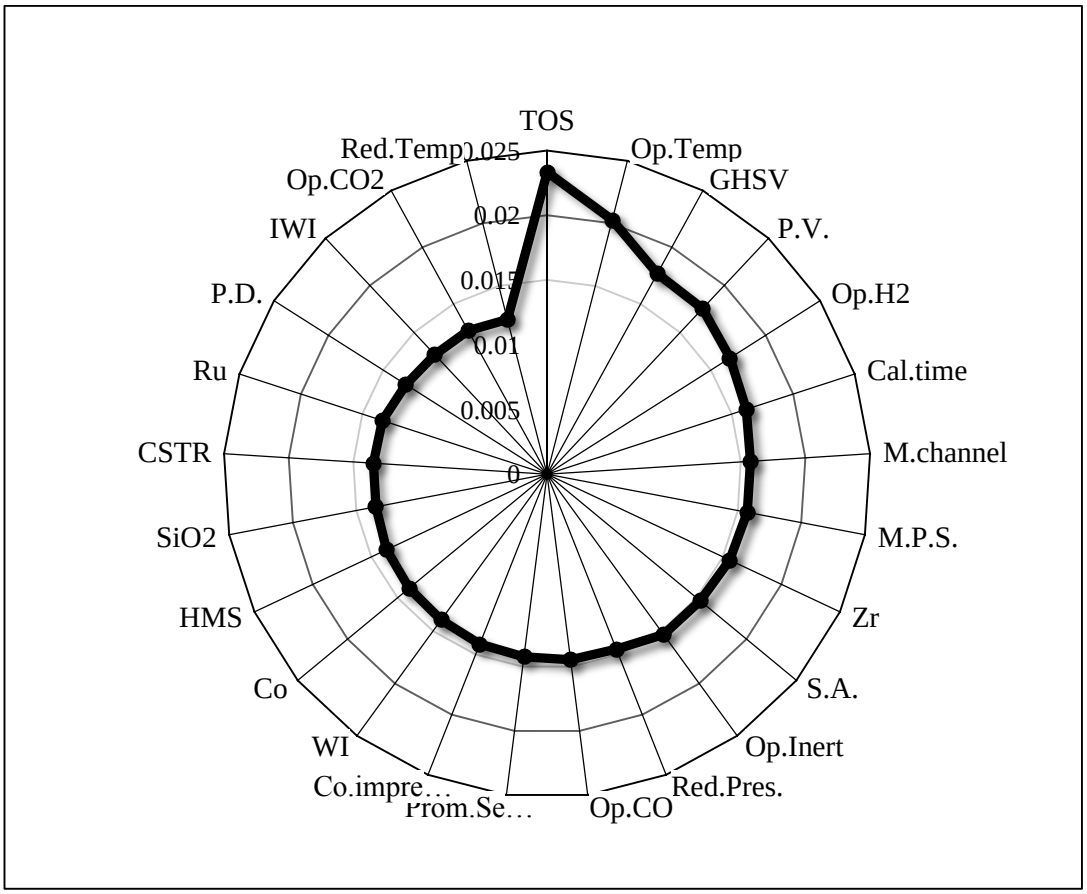

Figure 4.15. Relative importance of the attributes according to the Co-based training set.

Figure 4.15 shows the relative importance of the top 25 attributes out of 90 attributes. It was observed that time on stream, operating temperature, and GHSV were the most significant attributes in the model. This indicated that the most influential attributes for FTS on Co-based catalysts are operating conditions. Along with these attributes, pore volume was determined as another significant attribute in the model. This result supported the findings of the model constructed by the complete dataset.

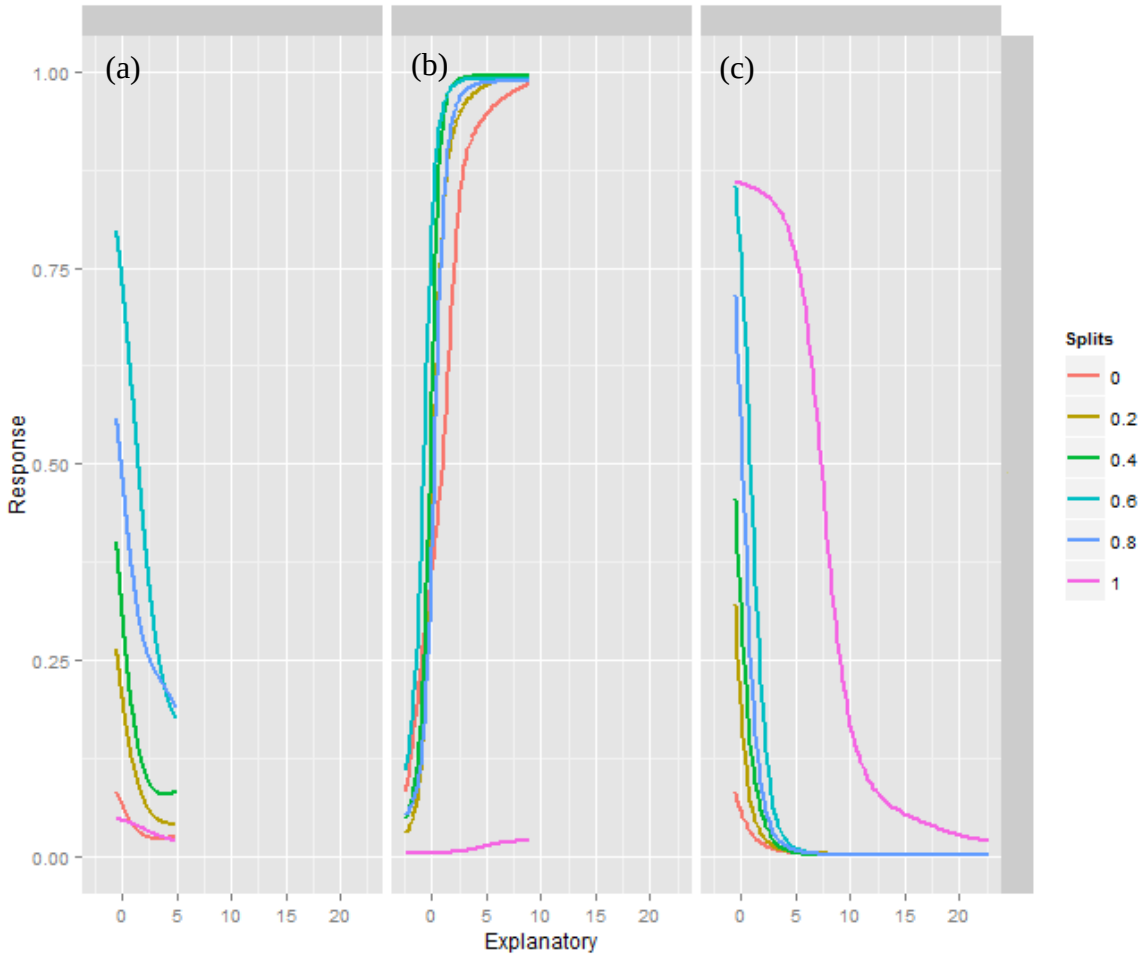

Figure 4.16. Sensitivity analysis for (a) TOS, (b) operating temperature, (c) GHSV.

The sensitivity analysis for the Co-based catalysts for 3 most important variables was demonstrated in Figure 4.16. It was observed that time on stream inversely affects the CO conversion at all split values. This effect can be explained by deactivation of the catalyst with respect to time. It was noted that deactivation rate was slightly slower at the highest split value. The effect of temperature was found to be similar with the model constructed for

the complete dataset. An increase in temperature significantly enhanced the CO conversion throughout the complete range. What was different from the previous sensitivity analysis in Figure 4.10 was that CO conversion profiles in Figure 4.16 continuously increased, whereas the profiles in the former decreased in the middle split values after specific maxima. GHSV also adversely affected the CO conversion. This result was reasonable since the time reactants spend in the reactor volume decreases as GHSV increases. The PCA results presented in Figure 4.5b supported this conclusion. It was noted that decrease in CO conversion was slightly smoother at the highest split value, whereas the decrease was significantly sharper at other splits.

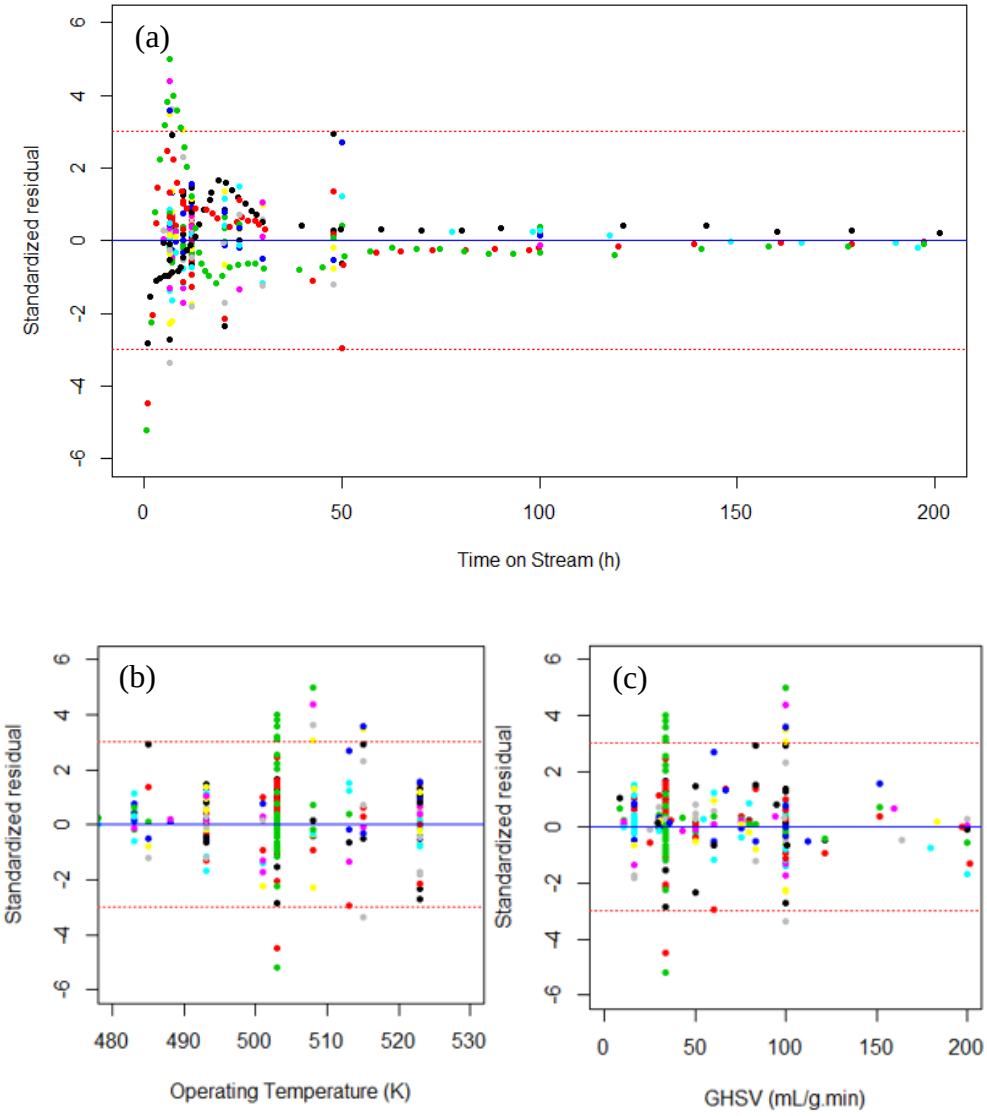

Figure 4.17. Standardized residuals vs. (a) time on stream (b) operating temperature (c) GHSV.

Figure 4.17 shows the results of the residual analysis of the important variables, namely time on stream (a), operating temperature (b), and GHSV (c). Although residuals were randomly distributed when plotted against operating temperature and GHSV, their distribution was still problematic with respect to time on stream. When the figure was inspected by the experimental runs, significant patterns were observed on each side of the mean. Consequently, this model was unable to model the effect of time on stream as well.

4.5.2. Fe-based Subset

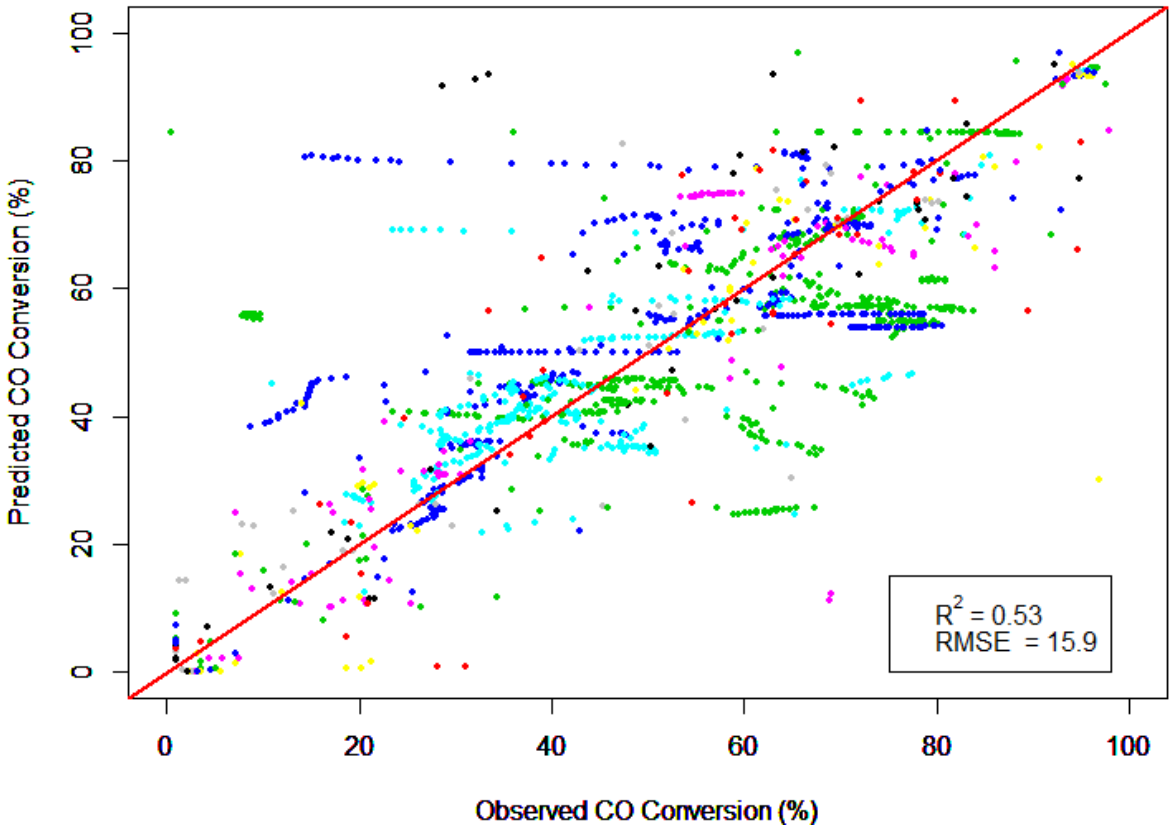

Figure 4.18. Predicted CO conversion vs. observed CO conversion for Fe-based training set.

The Fe-based subset consisted of 1900 instances in a 53 dimensional subspace. The best fitting ANN topology for Fe containing catalysts was obtained with 64 neurons and a bias unit in a single hidden layer. Logistic sigmoid was used as the activation function and standard backpropagation algorithm was used for learning. The optimum tuning for the

parameters were set at 0.05 for learning rate, 10^{-9} for absolute tolerance, and 100 for maximum number of iterations.

Figure 4.18 shows the prediction of the best fitting model by 10-fold cross-validation against their actual values. A significant improvement was observed with the Fe specific model when compared with the complete set model, regarding the increased R^2 value. Accordingly, it was concluded that splitting the complete database into two with respect to the base metal of the catalyst improved the accuracy for both of the subsets. Still, some of the experimental runs in this model were predicted to be constant throughout the course of reaction.

The accuracy of the model on the testing set was demonstrated in Figure 4.19. It was observed that R^2 and RMSE values had again improved compared to the model with the complete dataset. Still, TOS dependent data was poorly estimated. Although the predictions were reasonably accurate, the model failed to perceive any changes in time.

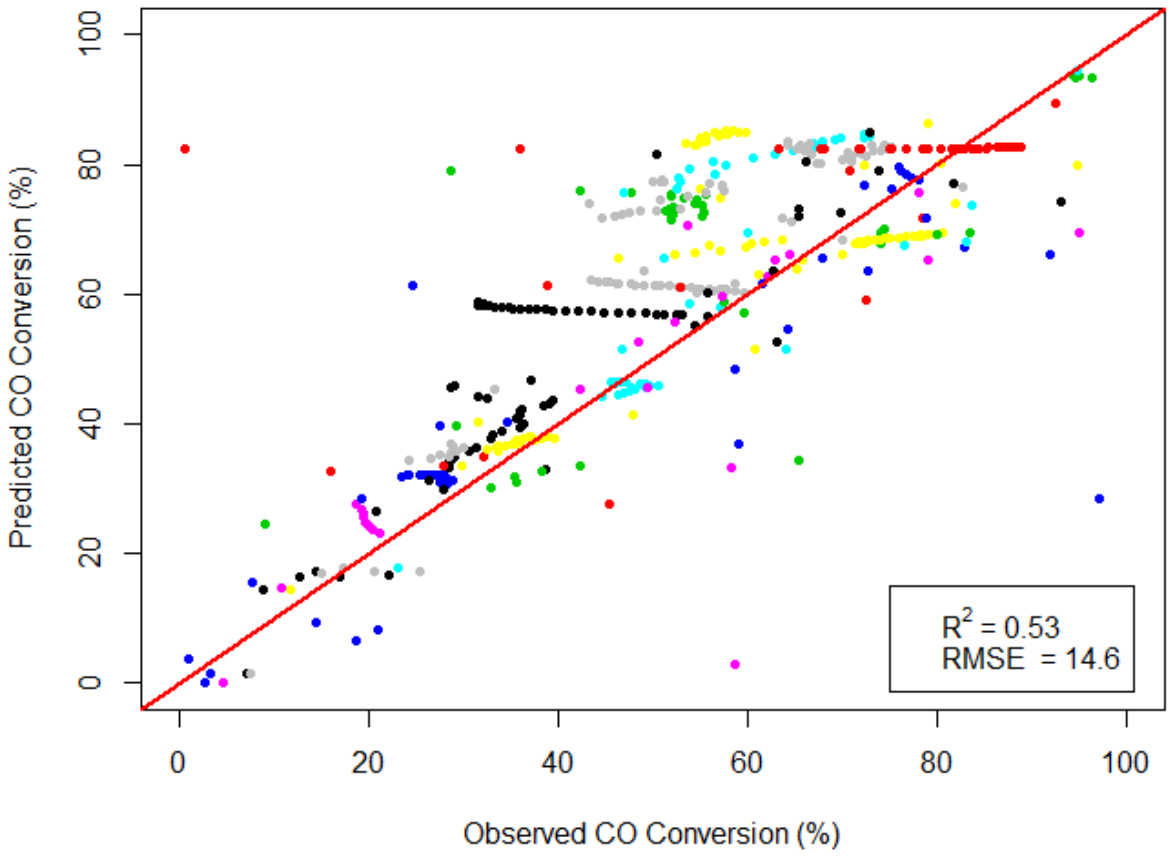

Figure 4.19. Predicted CO conversion vs. observed CO conversion for Fe-based testing set.

Comparing the combined performances of the two models constructed by using the subsets and the model constructed by the complete dataset, it was concluded that splitting the database with respect to the base metal improved the predictive power. In other words, more reliable estimations could be made by using the sub-models.

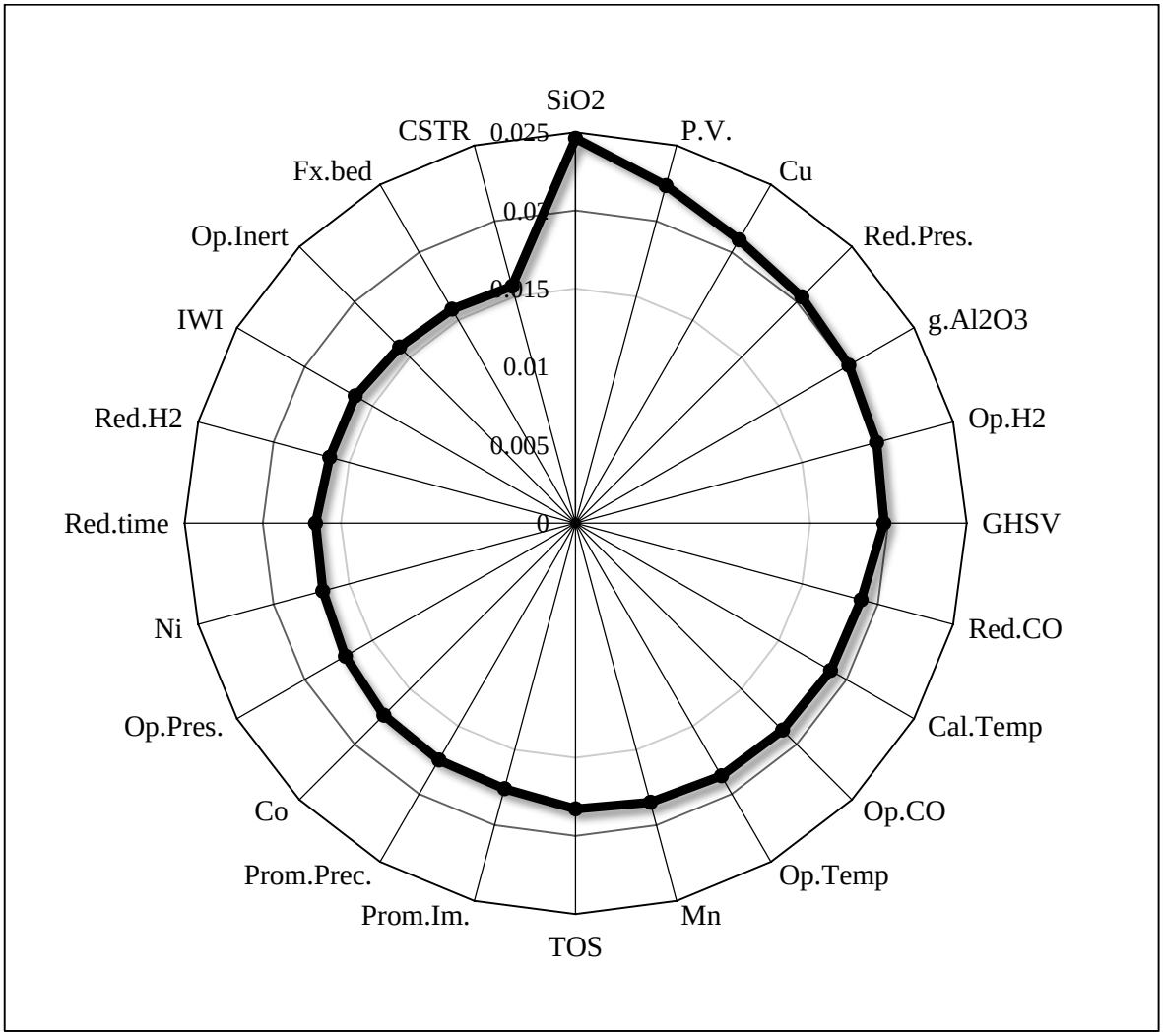

Figure 4.20. Relative importance of the attributes according to the Fe-based training set.

Figure 4.20 shows 25 of the most important attributes out of 53 attributes used to construct Fe-based ANN model. SiO_2 appeared to be the most important attribute for the Fe-based subset. This outcome was feasible since the performance of unsupported Fe-based catalysts are significantly improved by precipitating with small amounts support such as SiO_2 during preparation. Zhang *et al*. reported that addition of SiO2 to unsupported Fe-based

catalyst improved the C_2 – C_4 yield by improving selectivity while compromising CO conversion [209]. It should be noted that another frequently used support, γ-Al_2O_3 was also designated as a crucial attribute. Pore volume and Cu promoters were other important variables while constructing the model. It was noticed that unlike Co-based subset, the effects of catalyst design variables were much more dominant than operating variables in Fe-based subset.

The sensitivity analysis for the top three important attributes, namely SiO_2 loading, pore volume, and Cu loading, was shown in Figure 4.21. It was observed that promotion of Cu increases the CO conversion at all split values. This result was expected since the literature is in agreement on the promotional effect of Cu [6]. On the other hand, addition of SiO_2 on unsupported Fe catalysts improves performance up to a certain level. Therefore an optimization is required for SiO_2 loading when one deals with SiO_2 addition. Lastly, pore volume significantly enhanced CO conversion at all split values, which was found repeatedly in previous sections.

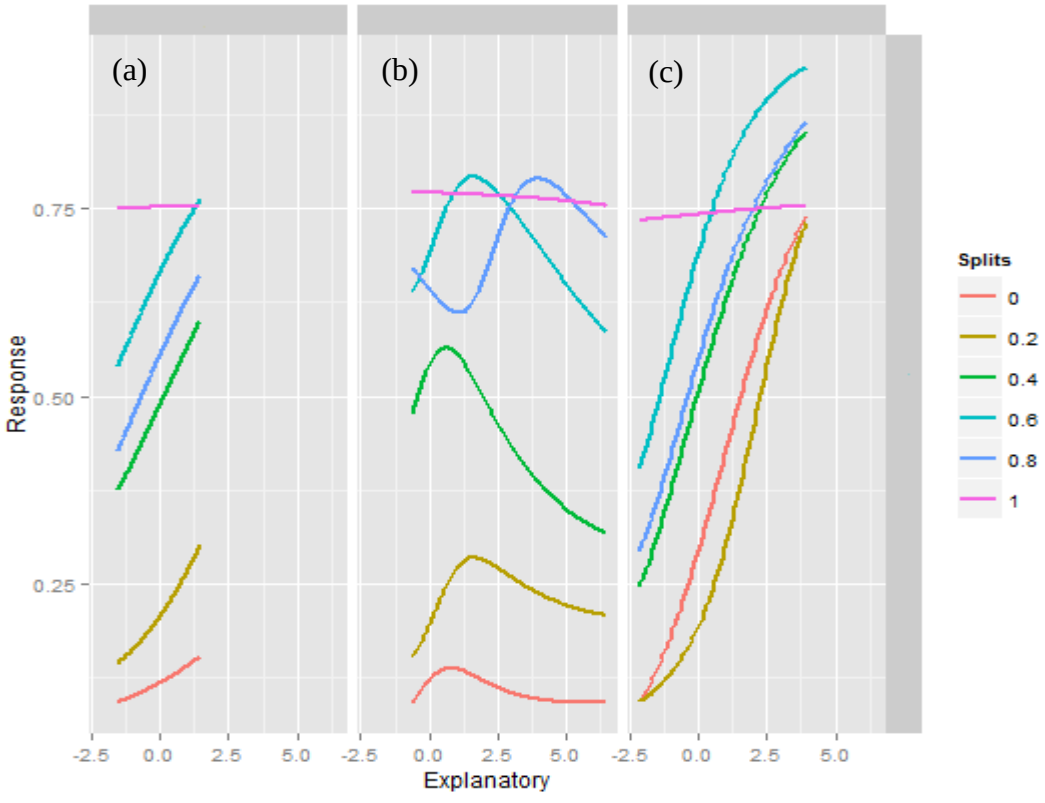

Figure 4.21. Sensitivity analysis for (a) Cu loading, (b) SiO_2 loading, (c) pore volume.

Figure 4.22 shows the residual analysis for the three important attributes. The standardized residuals were randomly distributed around the mean with respect to all three attributes. This indicated that the model could capture the behavior of the important attributes. However, the trend observed in previous models was revealed itself again with the residuals with respect to time on stream, as shown in Figure 4.23.

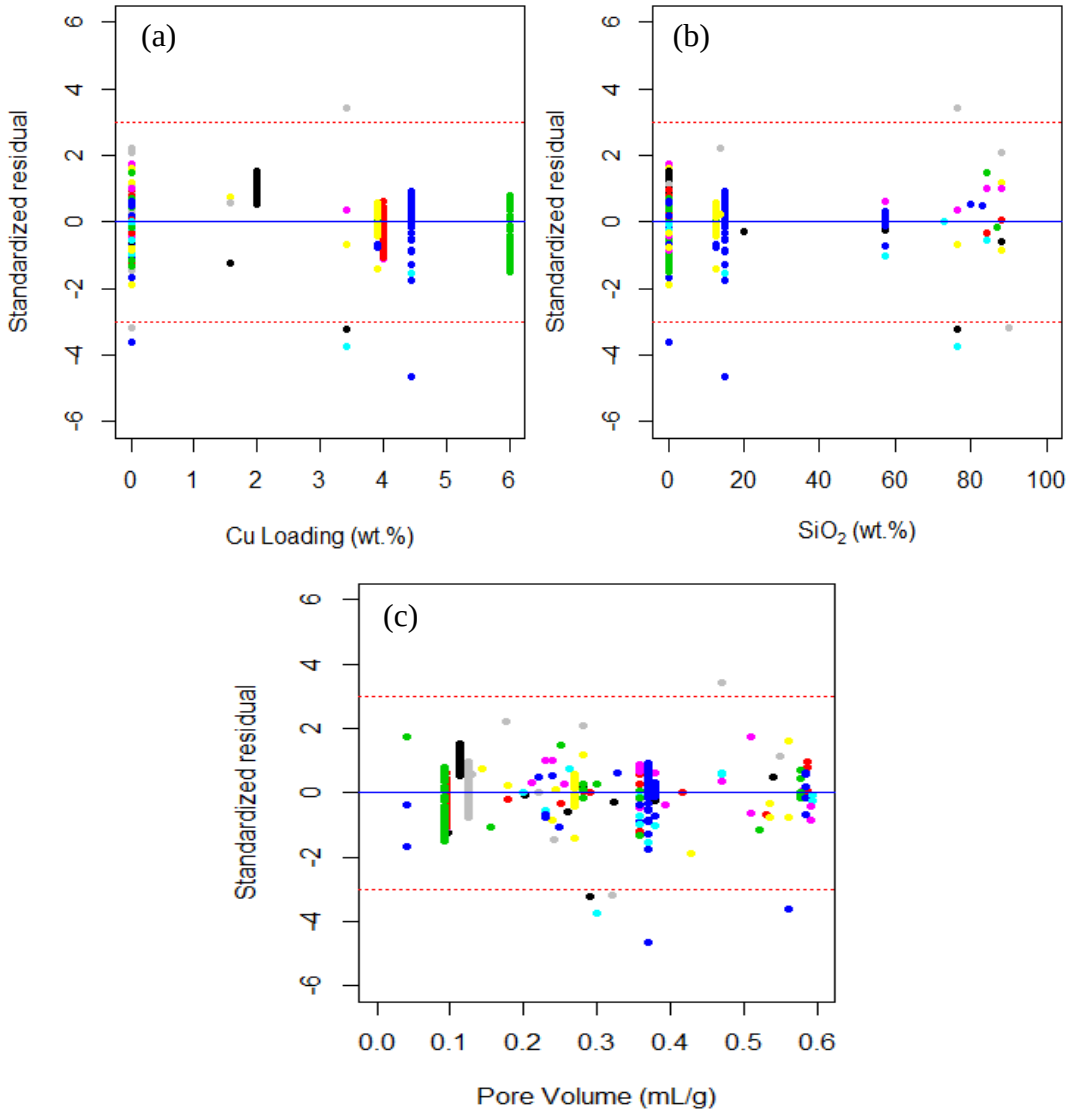

Figure 4.22. Standardized residuals vs. (a) Cu loading (b) SiO_2 loading (c) pore volume.

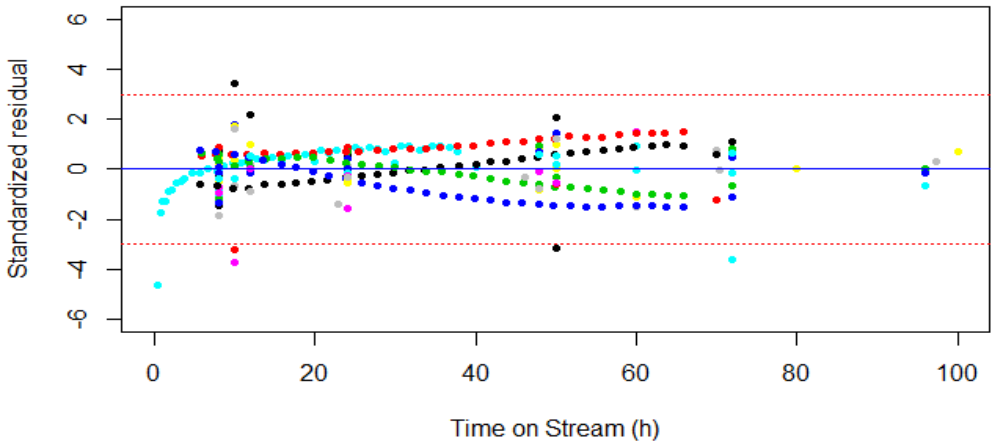

Figure 4.23. Standardized residuals vs. time on stream.

4.5.3. Low Temperature Subset

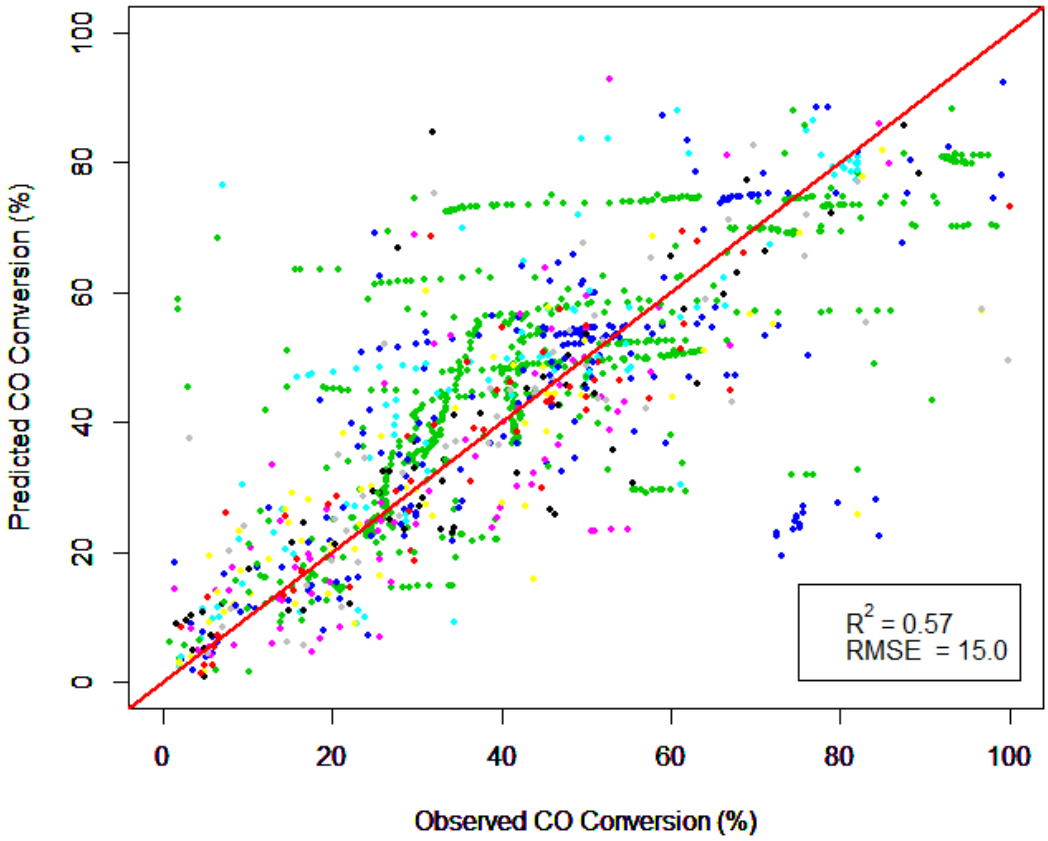

Figure 4.24. Predicted CO conversion vs. observed CO conversion for LT training set.

Low temperature (LT) subset consisted of 2103 instances of which operating temperatures were below 520 K. Dimensionality of LT subset was reduced to 43 attributes

after applying the method described in Figure 3.13. The best fitting network was obtained with 50 neurons in a single hidden layer, activated by logistic sigmoid function, and trained by standard backpropagation algorithm. The learning rate, absolute tolerance, and maximum number of iterations were set as 0.075, 10^{-9}, and 300, respectively to obtain the optimum network.

Figure 4.24 shows 10-fold cross validation predictions for the LT training set. A significant increase was observed in R^2 value compared to the cross-validation result of the complete dataset. Some of the TOS dependent experimental runs were estimated with improved accuracy. However, it can be concluded that the model failed to grasp the effect of time sufficiently, regarding the constant predictions for the experimental runs marked with green.

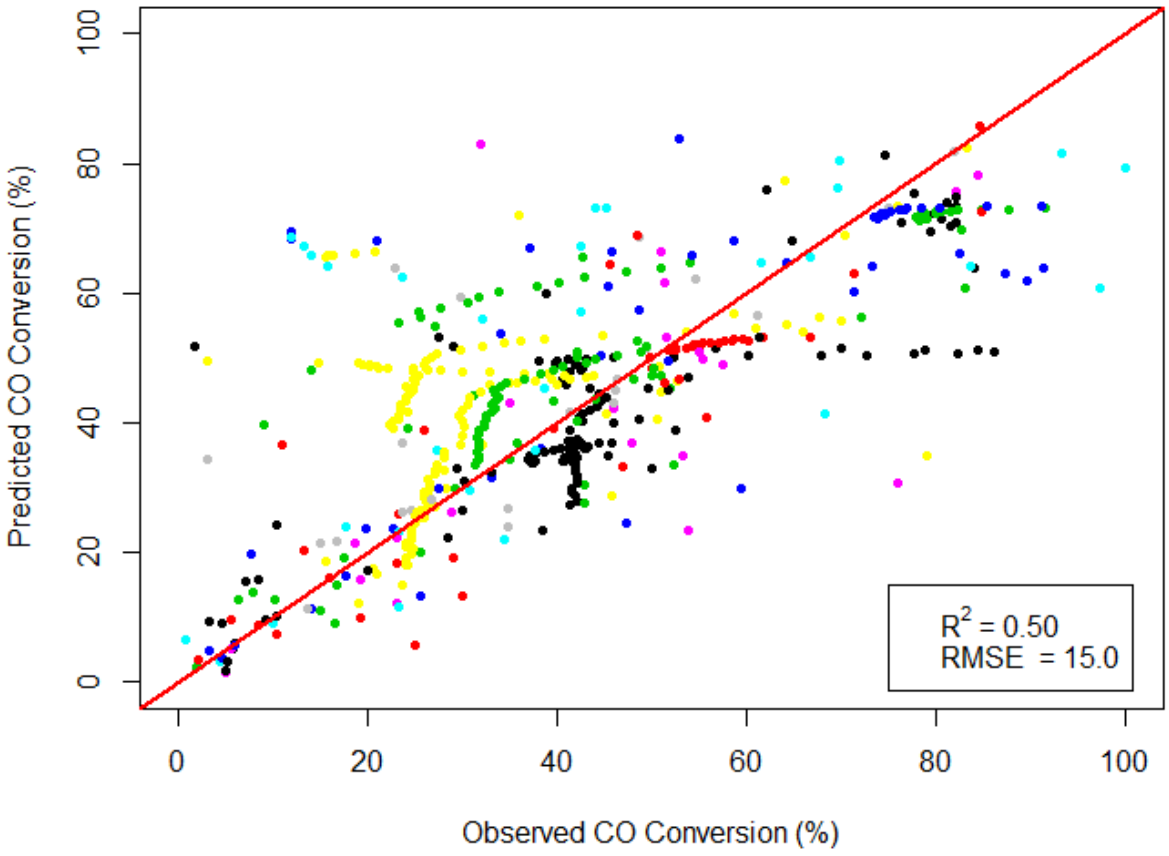

Figure 4.25. Predicted CO conversion vs. observed CO conversion for LT testing set.

Figure 4.25 shows the accuracy of the model on the previously separated testing set. It was observed that TOS dependent experimental runs were predicted with much higher

accuracy compared to the first ANN model discussed in Section 4.4. Accordingly, R^2 and RMSE values had improved significantly.

The importance analysis was carried out by the model constructed with the training set as shown in Figure 4.26. The first 25 attributes among a total of 43 were demonstrated. Time on stream and operating temperature were designated as the most important operating variables. Also, using γ-Al_2O_3 appeared to be a significant component among the catalyst design variables. However, this finding was questionable since the model could be biased toward γ-Al_2O_3 due to frequent appearance in the complete dataset, as shown in Table 3.4. Calcination time and pore volume were other significant attributes in the model. Different than the previous models, a catalyst preparation method, incipient-to-wetness method, appeared as an important component of the model.

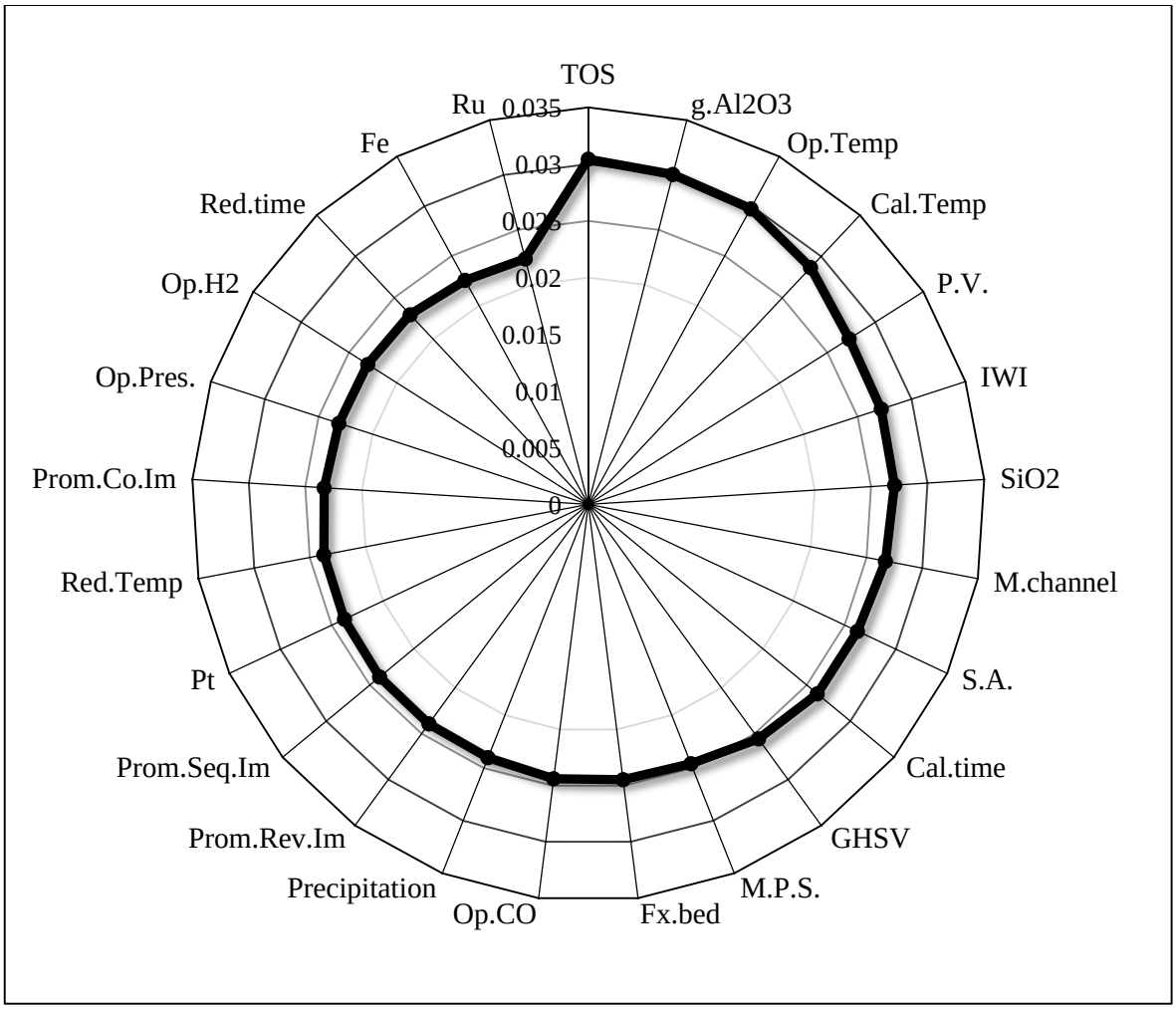

Figure 4.26. Relative importance of the attributes according to LT training set.

Figure 4.27 shows the sensitivity analysis on time on stream (a), operating temperature (b), and γ-Al_2O_3 loading (c). The effects of time on stream and operating temperature on CO conversion were similar to the results presented in Figure 4.16a and Figure 4.16b and thus, the same discussions were valid for the LT case. On the other hand, γ-Al_2O_3 loading appeared to be inversely affecting the CO conversion at most of the split values. The reason for this result can be explained by the fact that increasing the γ-Al_2O_3 loading implies decreasing the loading of the base metal and promoter, i.e. reducing the active sites available on the catalytic surface.

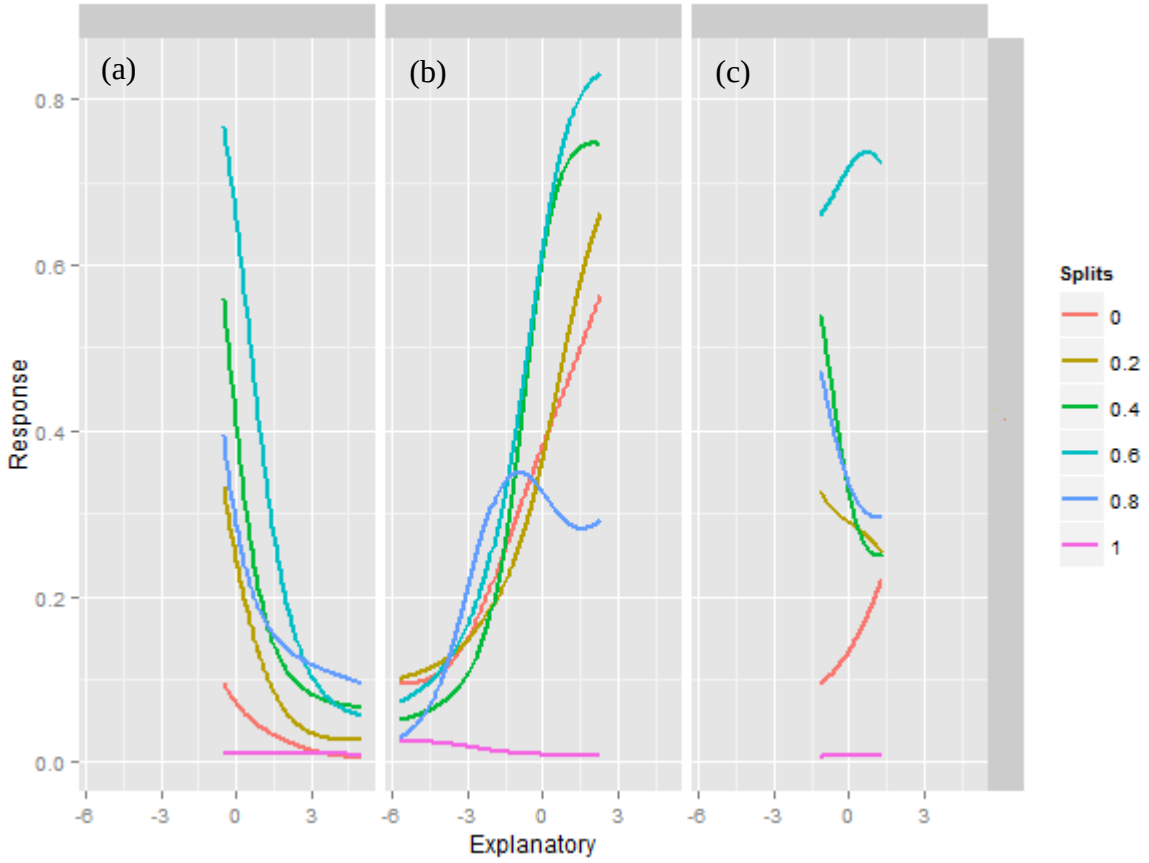

Figure 4.27. Sensitivity analysis for (a) TOS, (b) operating temperature, (c) γ-Al_2O_3 loading.

Figure 4.28 shows the residual analysis for time on stream, operating temperature and γ-Al_2O_3 loading. No outliers were detected within 99% confidence interval since all the residuals lied within the boundaries. However, a pattern was detected between the residuals with respect to time on stream and γ-Al_2O_3 loading as in the previous cases. It was observed

that the model tends to predict lower than the actual values at time on stream values below 100 h, whereas the prediction are above their actual values at the remaining interval. A similar but less obvious pattern was valid for γ-Al_2O_3 loading. The predictions were mostly higher at the interval between 60 – 80 wt.% interval, while they were below the mean between 80 – 100 wt.%. Although these two attributes were designated to be essential, their effects were modelled with a significant bias.

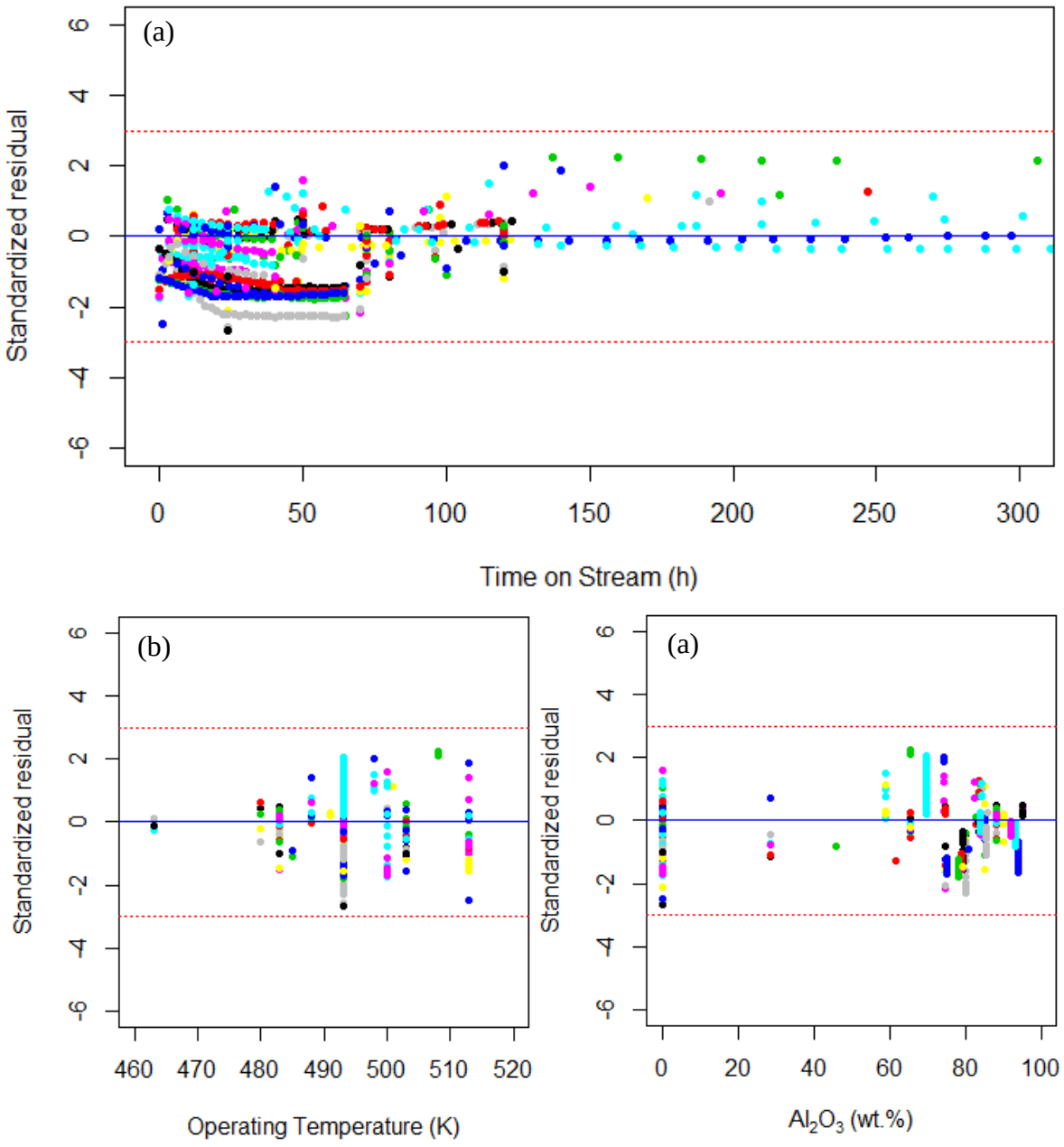

Figure 4.28. Standardized residuals vs. (a) TOS (b) operating temperature (c) Al_2O_3 loading.

4.5.4. High Temperature Subset

High temperature (HT) subset consisted of 1516 instances with operating temperatures above 520 K, and 37 attributes. The best fitting network consisted of 37 units and a bias in a single hidden layer, activated by logistic sigmoid function, and trained by standard backpropagation algorithm. The optimum values for learning rate, absolute tolerance, and maximum number of iterations were tuned as 0.075, 10^{-9}, and 75, respectively.

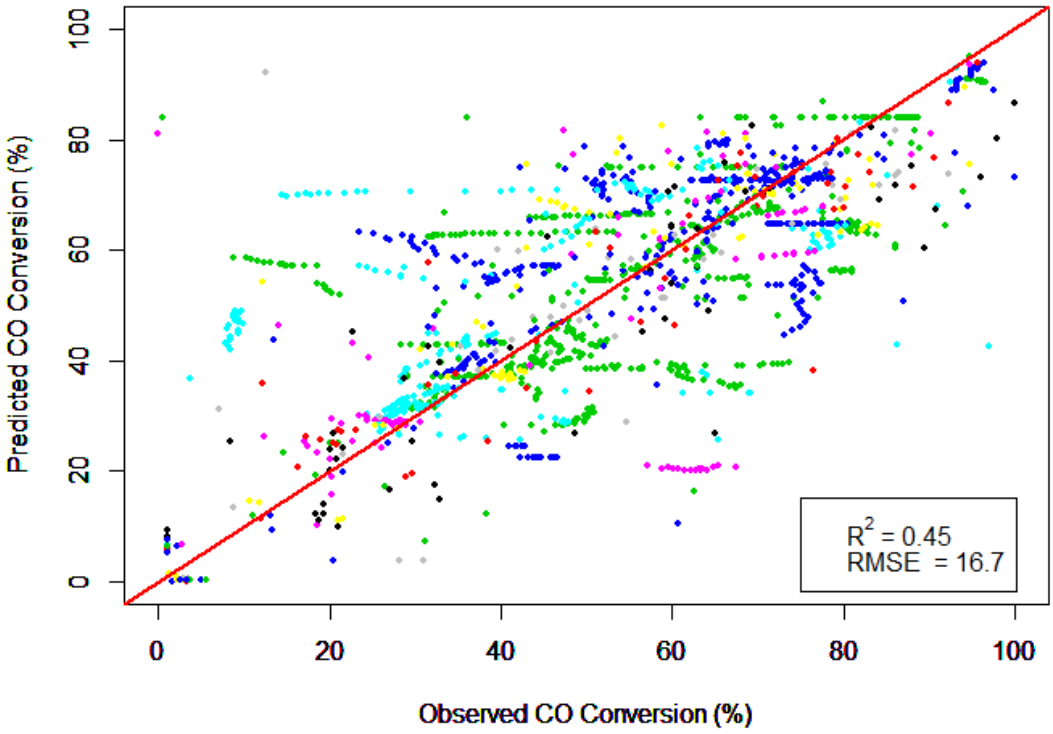

Figure 4.29. Predicted CO conversion vs. observed CO conversion for HT training set.

Figure 4.29 shows 10-fold cross validation predictions for the HT training set. Regarding the slightly increased RMSE; the HT model yielded worse predictions than the model constructed with the complete dataset, let alone giving improved results. Nevertheless, splitting the dataset with respect to the operating temperature still improved the overall predictive power because of high prediction accuracy of LT model. It was also observed that the model tended to shift the estimations towards the mean values, i.e. between 30 and 70%.

The accuracy of the model on the testing set was demonstrated in Figure 4.30. The improvement in R^2 and RMSE were hardly significant. However, considering the combined performance with the LT model, it was concluded that using the sub-models instead of the model trained by the complete dataset yields more accurate predictions.

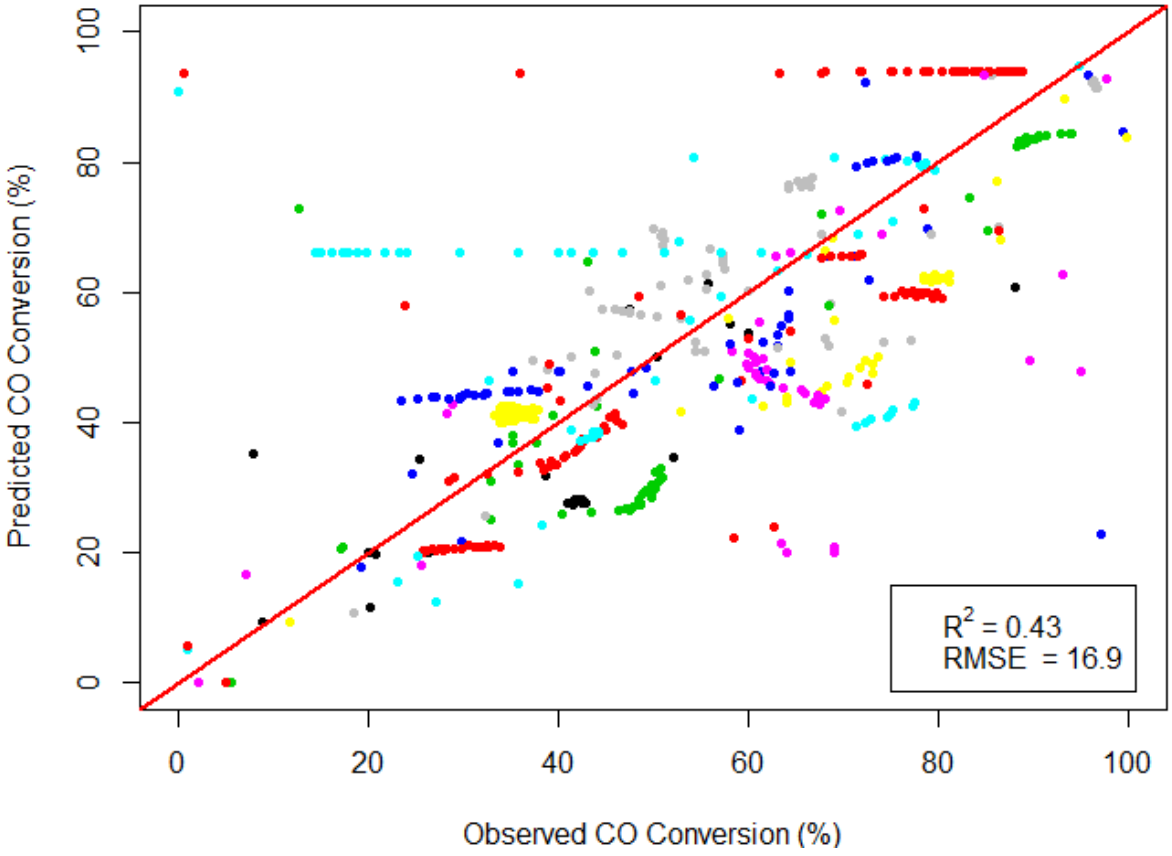

Figure 4.30. Predicted CO conversion vs. observed CO conversion for HT testing set.

Figure 4.31 shows the top 25 most important attributes among a total of 37. Although GHSV appeared as the most important constituent of the model, catalyst design variables such as SiO_2, Cu, K, and Mo loading were significant attributes as well. This result was reasonable since HTFT is strongly associated with Fe-based catalysts both in the literature and in the industrial applications; and these catalysts are mostly designed with the aforementioned promoters. This claim was supported by Fe being one of the most important attributes in Figure 4.31.

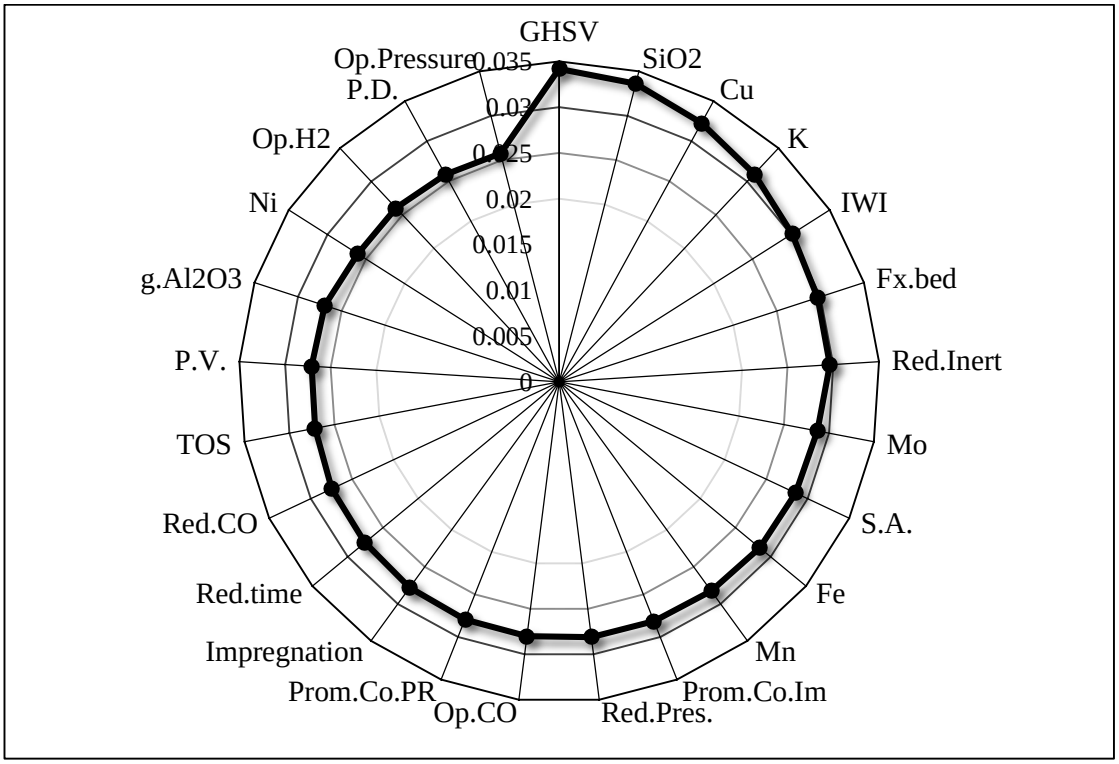

Figure 4.31. Relative importance of the attributes according to HT training set.

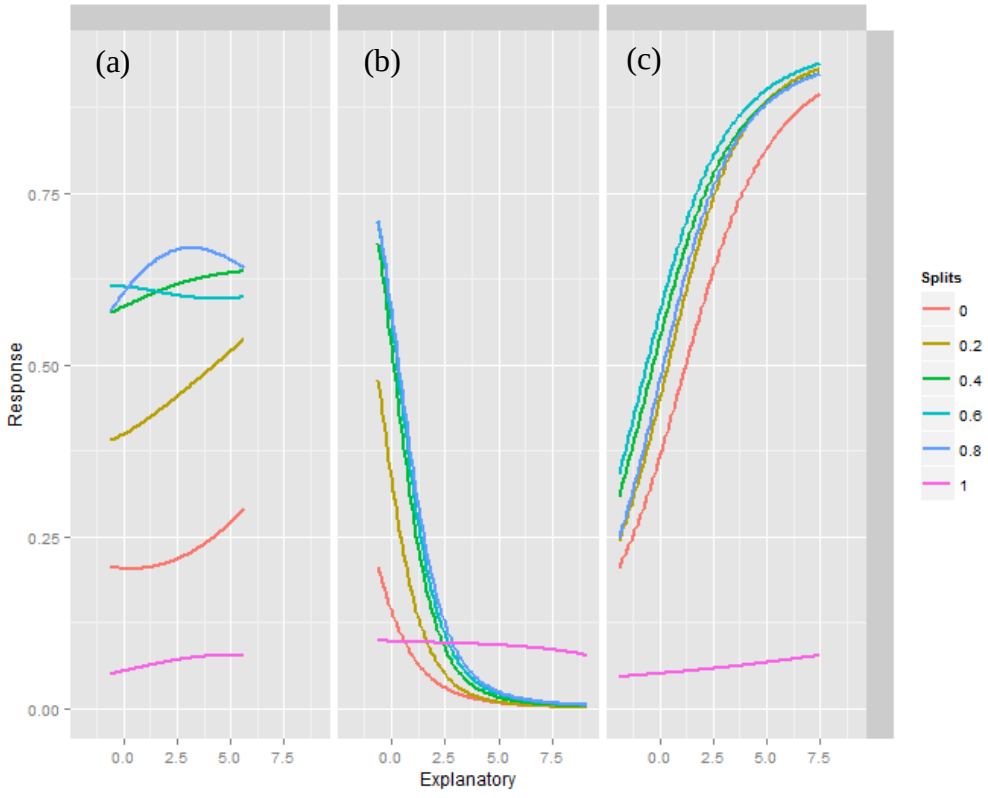

Figure 4.32. Sensitivity analysis for (a) SiO_2 loading, (b) GHSV, (c) Cu loading.

Sensitivity analysis for SiO_2 loading (a), GHSV (b), and Cu loading (c) were demonstrated in Figure 4.32. SiO_2 loading slightly increased the CO conversion whereas the effects of GHSV and Cu loading were much more apparent.

Figure 4.33 shows the residual analysis for the 3 important attributes and time on stream. A mild pattern was observed in GHSV in Figure 4.33b. The predictions tended to be lower than the observed values at GHSV values up to 100 mL/g.min, whereas they were higher at high GHSV values. Also, distinct patterns were observed in Figure 4.33d, indicating that the model failed to represent the effect of time on stream as expected.

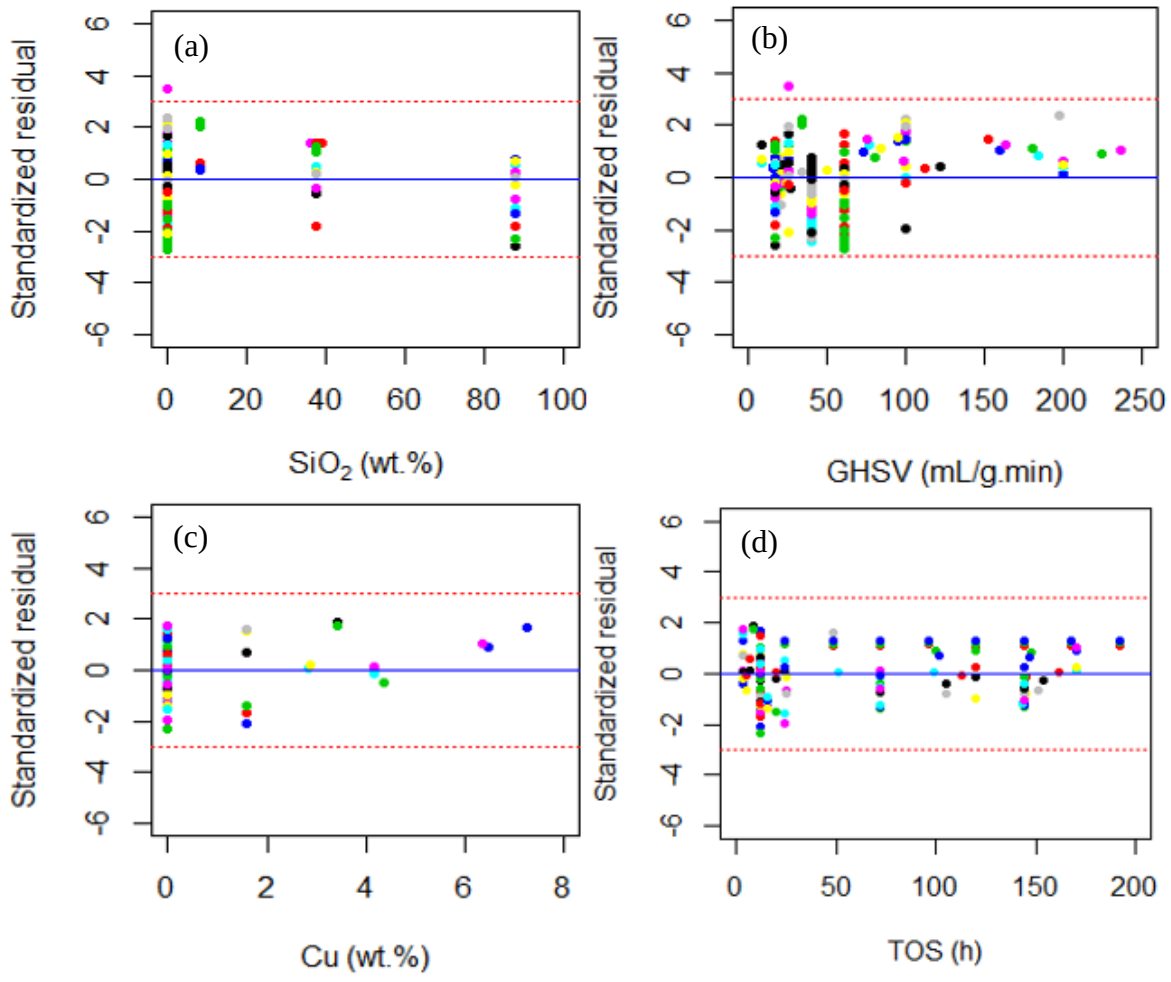

Figure 4.33. Standardized residuals vs. (a) SiO_2 loading (b) GHSV temperature (c) Cu loading, (d) TOS.

4.6. Decision Tree Prediction

Before constructing the model, the CO conversion values were discretized according to the intervals provided in Section 3.2.5. The best fitting decision tree model was constructed by using the 52 attributes used in the ANN model. The optimum tree was determined by trial-and-error method, and the misclassification error was employed as a measure of success. The minimum number of observations in the terminal nodes and the maximum depth of the tree were determined to be 3 and 15, respectively for the optimum model. Other parameters, complexity parameter and maximum number of competing splits were set to be 10^{-3} and 7, respectively.

Table 4.3 shows the confusion matrix of the training set obtained by 10-fold cross-validation. The misclassification error was computed to be 43.9%, indicating that the decision tree algorithm was unsuccessful to model the FTS database.

Table 4.3. 10-fold cross-validation accuracy of the decision tree model.

		Observed CO Conversion			
		Very Bad	Bad	Good	Very Good
Predicted CO Conversion	Very Bad	370	109	48	12
	Bad	308	874	259	85
	Good	70	301	387	62
	Very Good	18	31	45	89

Table 4.4. Testing set accuracy of the decision tree model.

		Observed CO Conversion			
		Very Bad	Bad	Good	Very Good
Predicted CO Conversion	Very Bad	122	80	52	18
	Bad	52	367	67	17
	Good	7	52	77	8
	Very Good	32	8	23	48

Table 4.4 shows the accuracy of the model on the testing set on a confusion matrix. The misclassification error for this set was computed to be 40.4%, slightly lower than the cross-validation error. Thus, it was concluded that the predictive power of the decision tree model was weak, and no further effort was spent on improving the model.

4.7. Random Forest Prediction for Entire Dataset

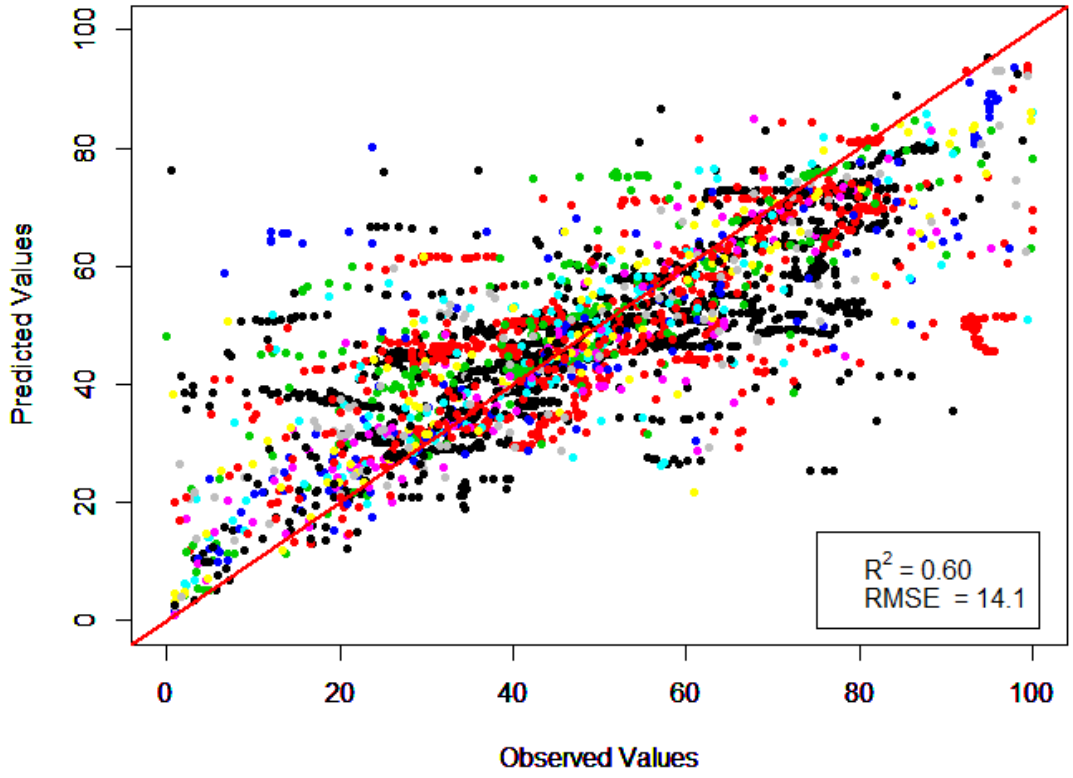

Figure 4.34. Predicted CO conversion vs. observed CO conversion for complete training set.

The best fitting random forest model was constructed by following the procedure described in Figure 3.13. The dataset consisted of 51 dimensions after eliminating the least important attributes. Each tree was trained with 17 randomly selected attributes. Instead of using all of the instances, they were sampled by bootstrapping for each tree. Each terminal node in the trees contained at least five instances. Also, the trees were grown to the maximum possible depth, i.e. no limitations were enforced on the size of the trees. Number of trees in a forest, denoted by “ntree” argument in “randomForest” function, was the only optimized parameter. A wide range from 10 to 1000 trees were grown in each forest to find the optimum

forest size. The forests with size larger than 140 trees yielded the lowest RMSE and no significant increase was observed at more crowded forests. Therefore the resultant model was constructed with 140 trees to keep the model as simple as possible.

Figure 4.34 shows the accuracy of the best fitting random forest model based on 10-fold cross-validation of the previously separated training set. It was observed that the accuracy of the model is significantly higher than of the model constructed by ANN, shown in Figure 4.7. The TOS dependent experimental runs were estimated with greater success, regarding that the instances with the same color were much more randomly distributed around the identity line. Nevertheless, a similar bias towards 40 – 60% interval was observed in these predictions as well.

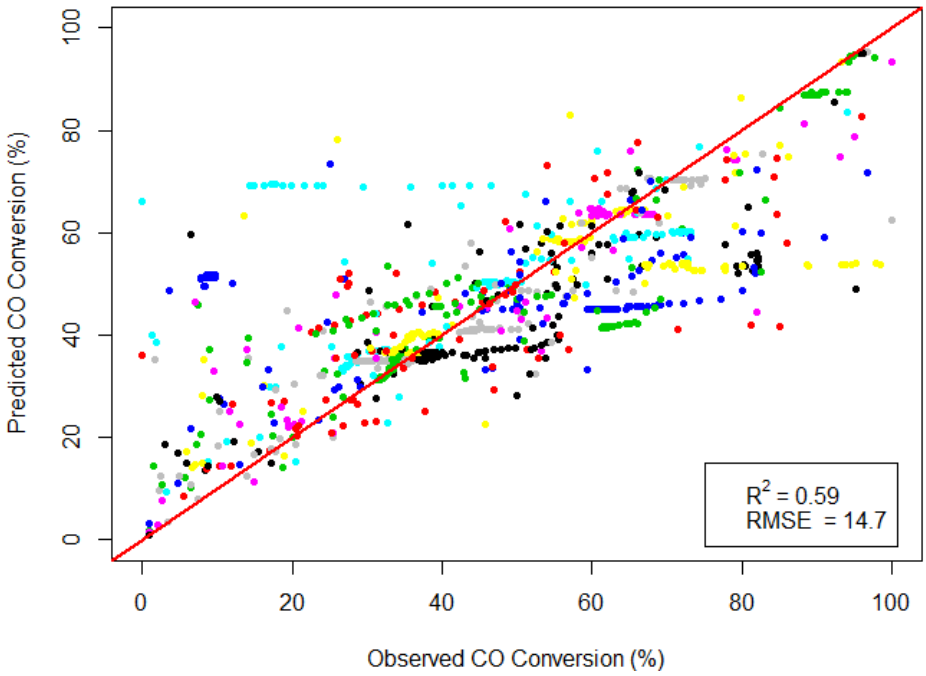

Figure 4.35. Predicted CO conversion vs. observed CO conversion for complete testing set.

Figure 4.35 shows the accuracy of the model on the previously separated testing set. Although some of the results were erroneous because of the bias of the model, the predictions were unequivocally more accurate than the results of the ANN model.

4.7.1. Importance Analysis

The relative importance of the attributes constituting the model was determined by computing the mean decrease in accuracy in their individual absence from the column space. Figure 4.36 shows the relative importance of the attributes on polar coordinates with decreasing order in clockwise. It was observed that the operating conditions have a crucial role on CO conversion since operating temperature, GHSV, and operating pressure were found to be among the most important attributes. Additionally, all the physical properties, i.e. pore diameter (P.D.), metal particle size (M.P.S.), surface area (S.A.), and pore volume (P.V.) were designated as significant attributes.

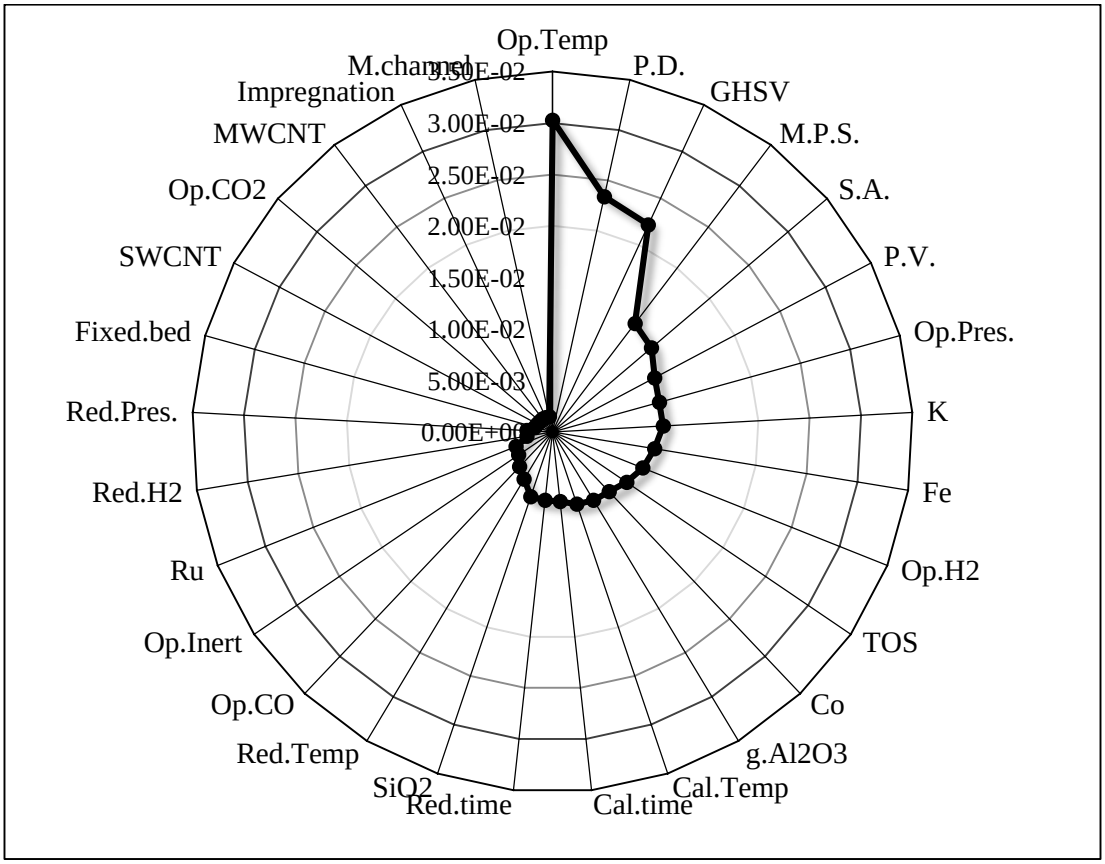

Figure 4.36. Relative importance of the attributes according to the complete training set.

It can be asserted that the importance analysis by random forest model is more reliable than that by ANN model, since its predictive power on unseen instances was significantly greater. However, the analyses attained from either model were in consistence with each other, albeit their priority for the more important attribute had slightly differed.

4.7.2. Residual Analysis

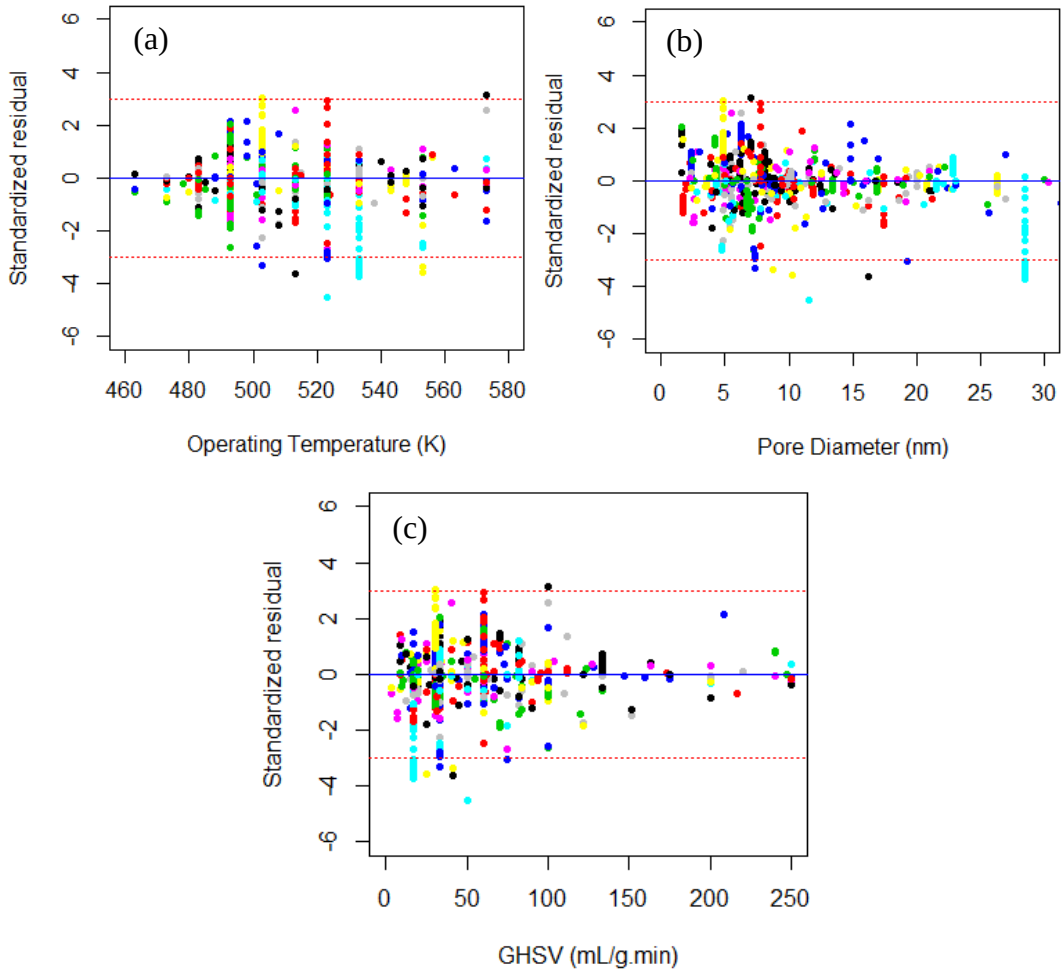

Figure 4.37. Standardized residuals vs. (a) operating temperature (b) pore diameter (c) GHSV.

Figure 4.37 shows the standardized residuals between the predictions and their corresponding actual values with respect to three of the most important attributes, namely operating temperature (a), pore diameter (b), and GHSV. It was observed that the residuals were well distributed around zero and no patterns were detected with respect to these attributes. This indicated that the model could capture the behavior of these attributes well. Only a few outliers were observed in 99% confidence interval.

Figure 4.38 shows the standardized residuals with respect to time on stream, which was the problematic attribute in the ANN model. Unfortunately the trends were also

observed in random forest analysis; TOS dependent experimental runs were poorly modelled as it is evident from the behavior of the instances belonging to same experiment at different time on stream (with same colors)

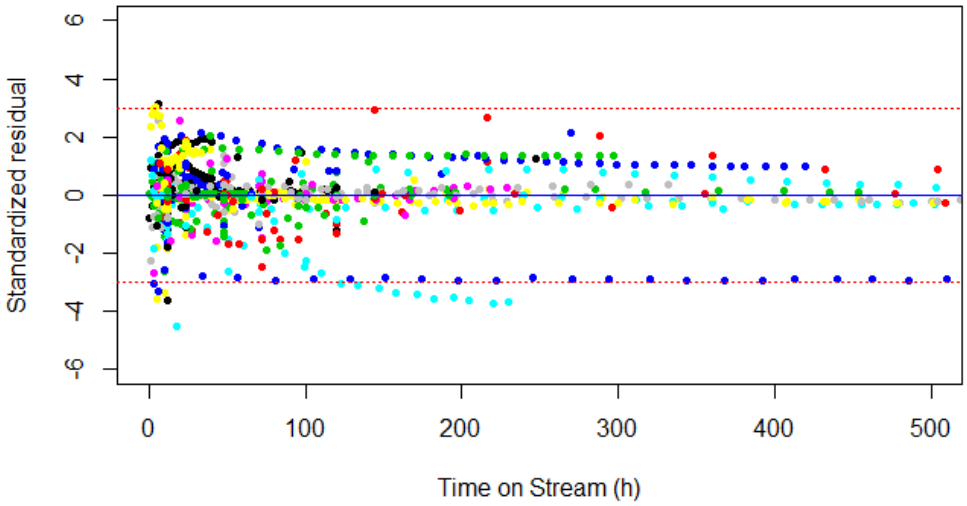

Figure 4.38. Standardized residuals vs. time on stream.

4.8. Random Forest Prediction for Individual Subsets

The aforementioned four subsets, namely Co-based, Fe-based, LT, and HT subsets were modelled by random forest algorithm as well.

4.8.1. Co-based Subset

The dimensionality of the Co-based subset was reduced to 57 attributes after applying the preprocessing steps depicted in Figure 3.13. Each tree in the forest was trained by 19 randomly selected attributes. The optimum forest size for modelling this subset was determined to be 160, after trying the interval given in Section 4.7.

Figure 4.39 shows the predictions for 10-fold cross validation obtained by the optimum random forest model. A slight improvement was observed in accuracy compared to the random forest model constructed with the complete dataset. Also, the model yielded more accurate predictions than the ANN model constructed by the same subset, since both R^2 and RMSE values had improved. The model appeared to be biased towards the interval 20 – 60%

CO conversion, but it was more sensitive to the TOS dependent experimental runs than its ANN counterpart.

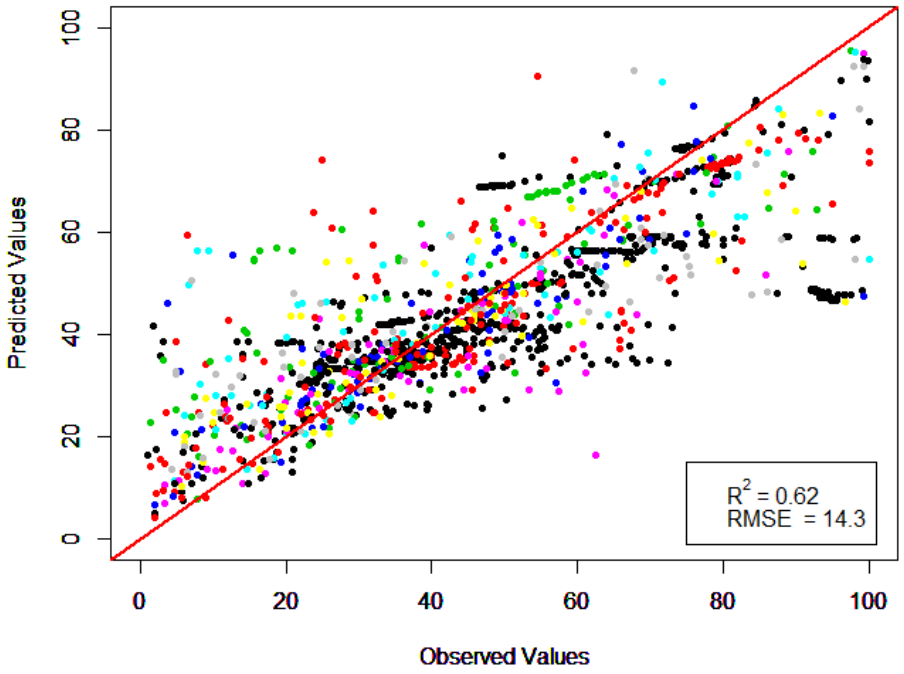

Figure 4.39. Predicted CO conversion vs. CO conversion for Co-based training set.

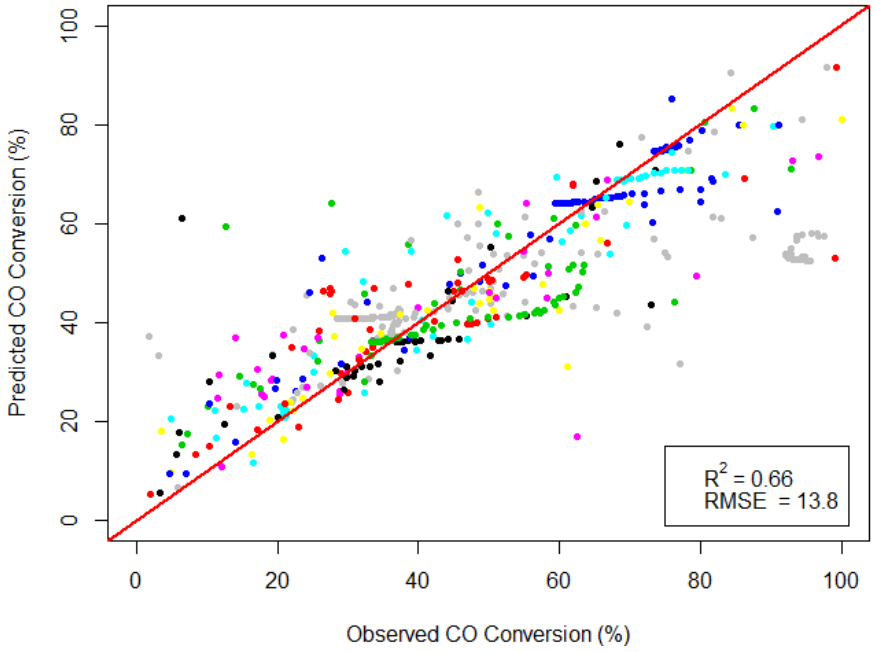

Figure 4.40. Predicted CO conversion vs. observed CO conversion for Co-based testing set.

The testing set was introduced to the model to investigate the validity of the model. Figure 4.40 shows the predictions of the model with respect to their actual values. It was observed that the accuracy of the model was significantly greater than the corresponding ANN model regarding the R^2 and RMSE values.

Figure 4.41 shows 25 of the most important attributes among a total of 57. It was observed that three operational attributes, namely operating temperature, GHSV, and hydrogen composition in the feed stream, ranked the top among other important attributes. This result was consistent with the ANN results presented in Figure 4.15, which revealed that operating conditions were more important than catalyst design variables. Random forest model also indicated that SiO_2 was an important attribute, yet this result might stem from the fact that most of the Co-based catalysts in the database were prepared on SiO_2 supports, which probably created a bias.

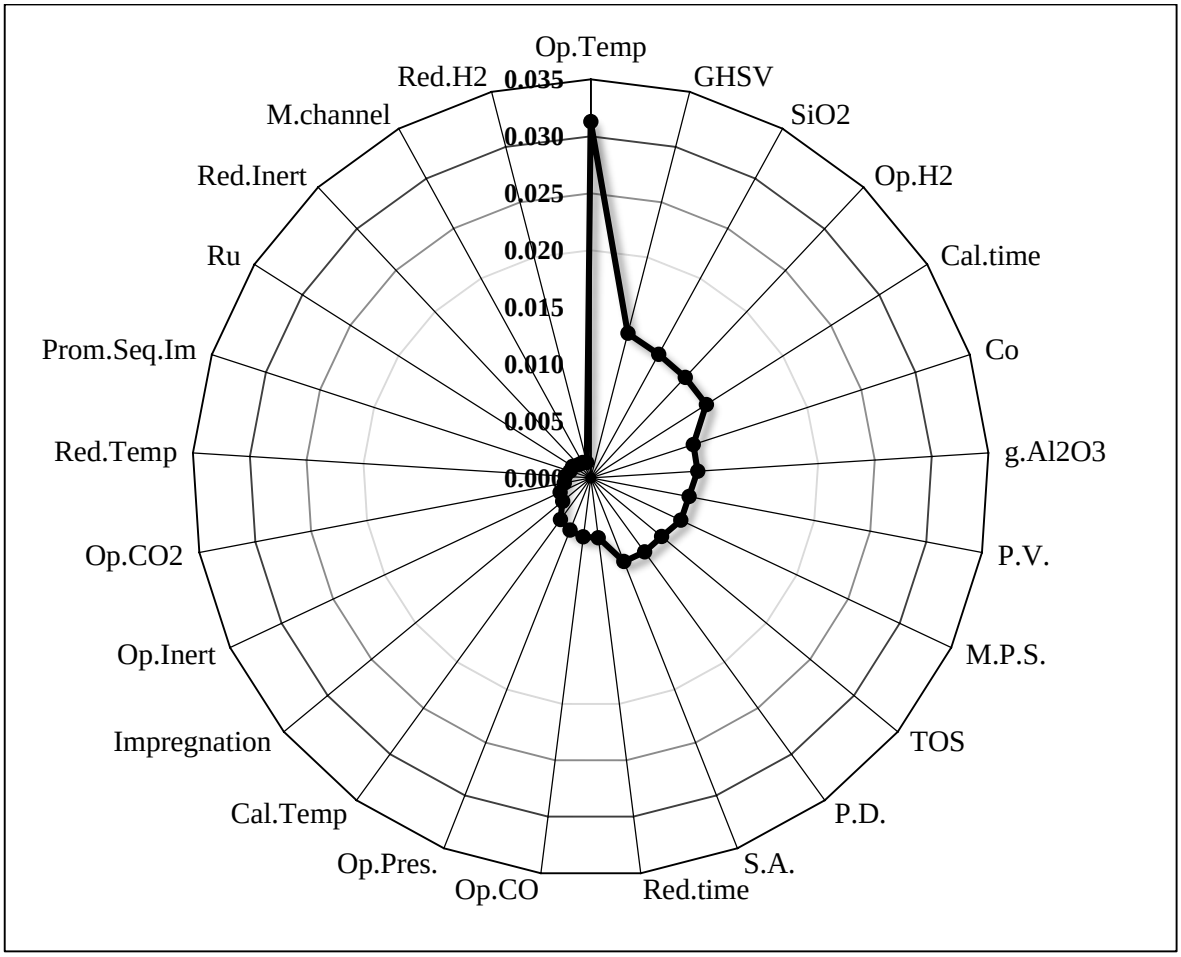

Figure 4.41. Relative importance of the attributes according to the Co-based training set.

Figure 4.42 shows the standardized residuals with respect to operating temperature, GHSV, and SiO_2 loading. It was observed that the residual were well distributed around zero, indicating that these attributes were well modelled by random forest algorithm. Yet, a few outliers were detected outside of 99% confidence interval.

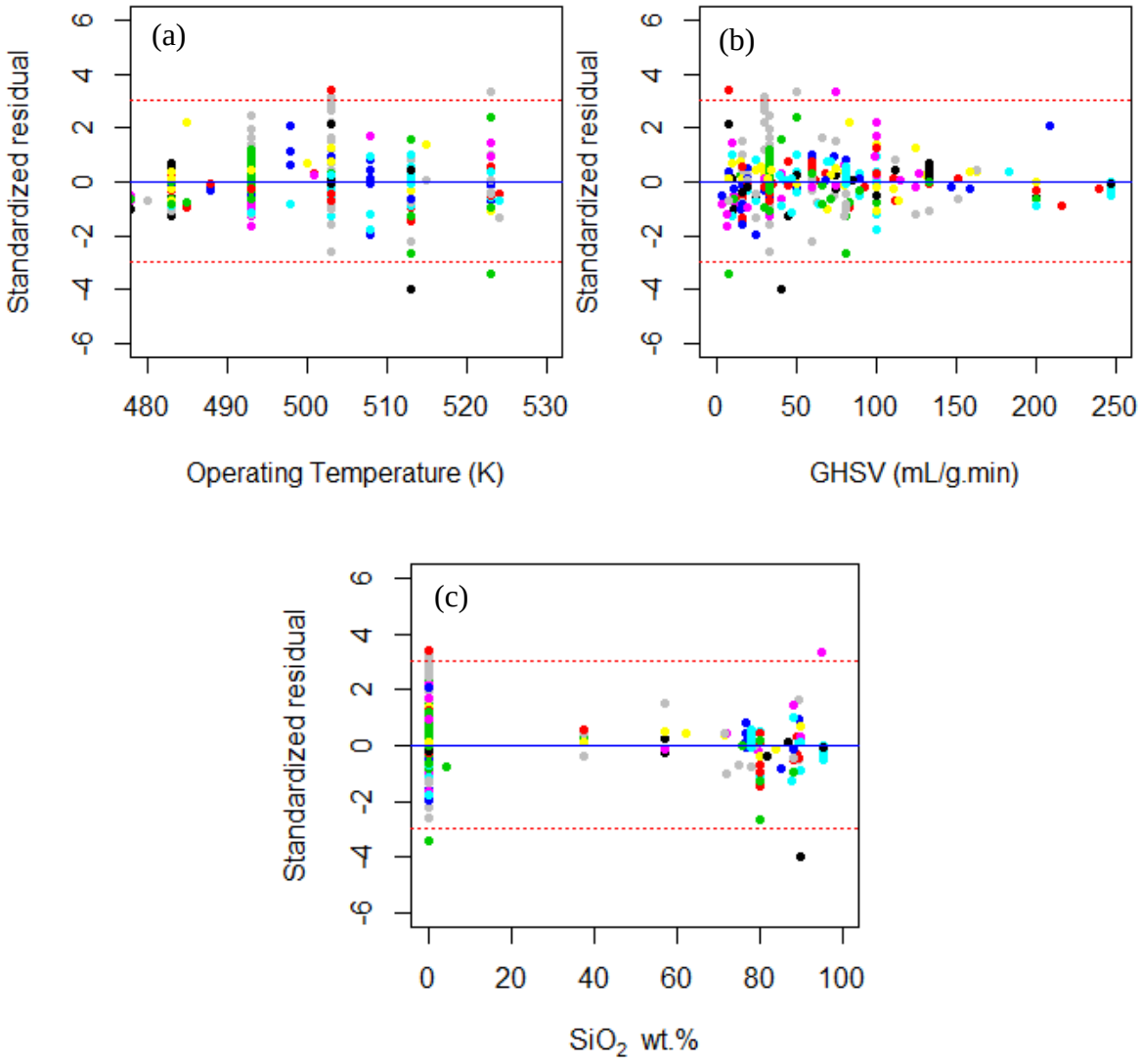

Figure 4.42. Standardized residuals vs. (a) operating temperature (b) GHSV (c) SiO_2 loading.

The same standardized residuals were also plotted against the time on stream in Figure 4.43, since it created the main problem in the model. It was observed that the residuals from the same experimental runs follow obvious patterns at time on stream values below 50 h.

However, they converged to zero and were distributed more randomly around zero after 50 h.

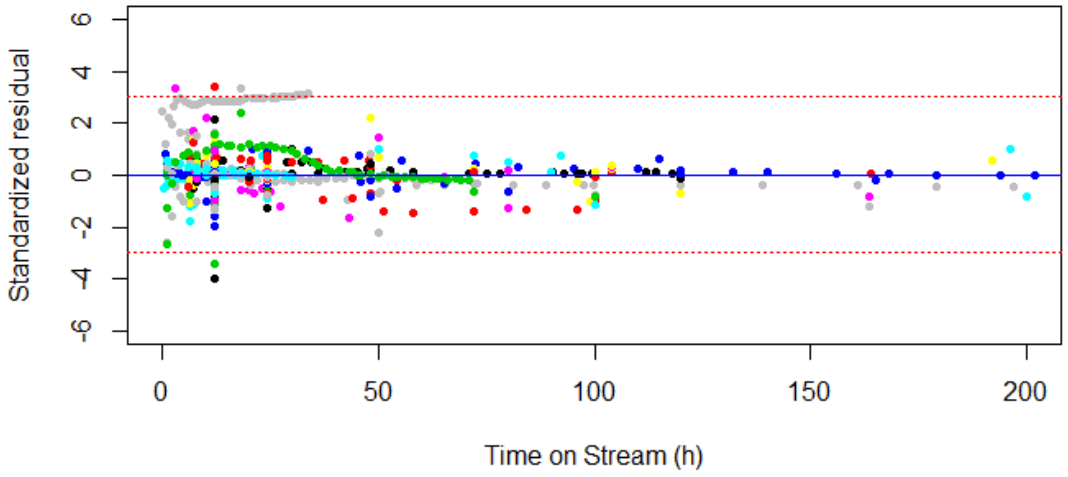

Figure 4.43. Standardized residuals vs. time on stream.

4.8.2. Fe-based Subset

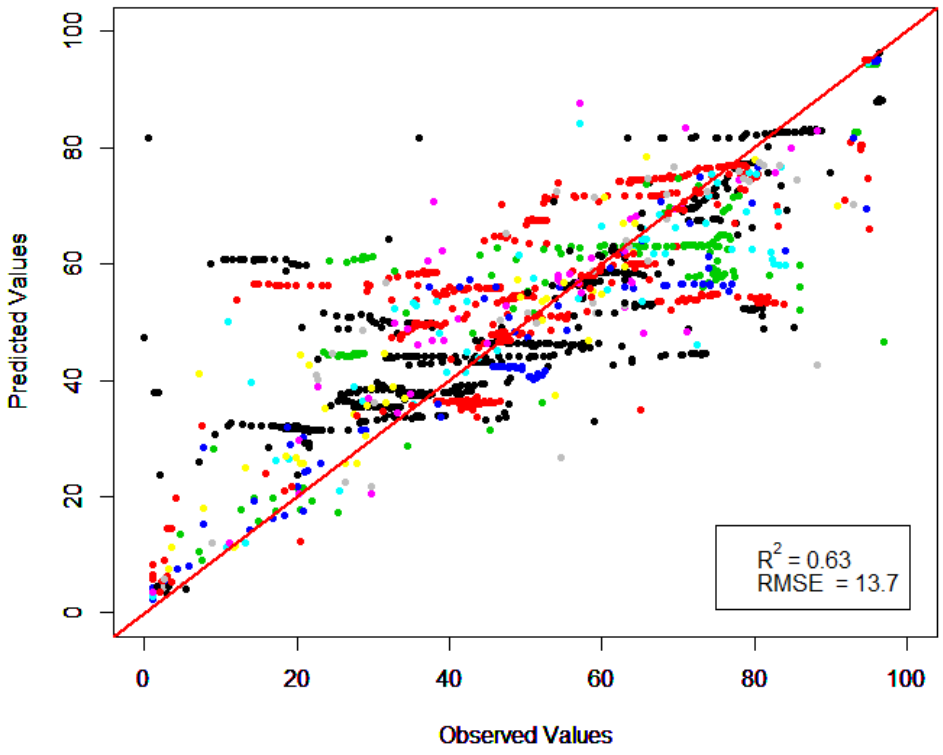

Figure 4.44. Predicted CO conversion vs. observed CO conversion for Fe-based training set.

The dimensionality of the Fe-based subset was reduced to 51, after applying the procedure described in Figure 3.13. Each tree was grown with 17 randomly selected attributes. The optimum number of trees in the forest was found to be 170.

Figure 4.44 shows the accuracy of the best fitting random forest model based on 10-fold cross-validation of the training set. It was observed that R^2 and RMSE values were slightly greater than those of the model constructed with the complete dataset; and significantly greater than the model trained by ANN with the same subset.

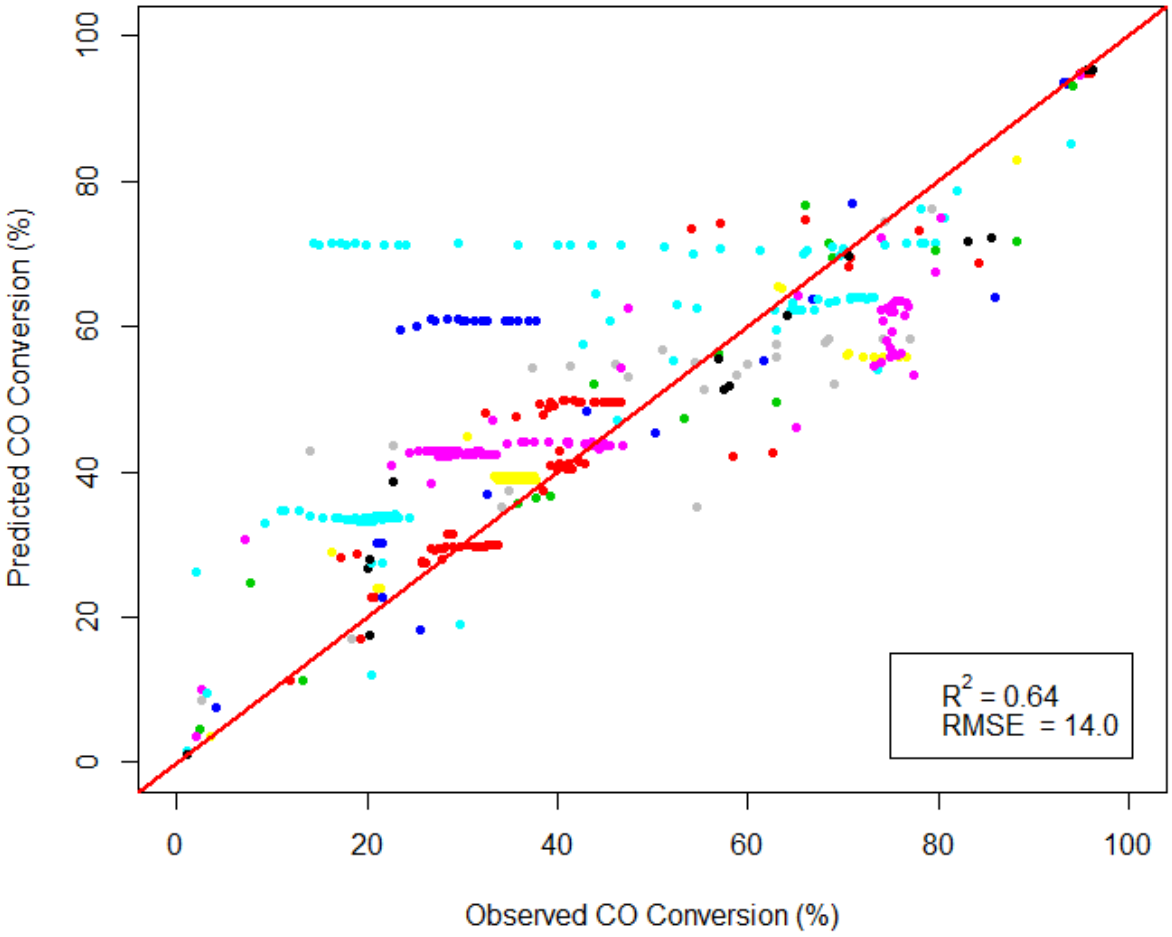

Figure 4.45. Predicted CO conversion vs. observed CO conversion for Fe-based testing set.

The predictions of the testing set by the optimum model was demonstrated in Figure 4.45. R^2 and RMSE values were observed to be increased significantly compared to the complete set. Accordingly, it was concluded that splitting the database with respect to the base metal to model with random forest algorithm enhances the predictive power of the

model. Also, it was observed that the performance of the random forest model is much greater than its ANN counterpart in terms of accuracy.

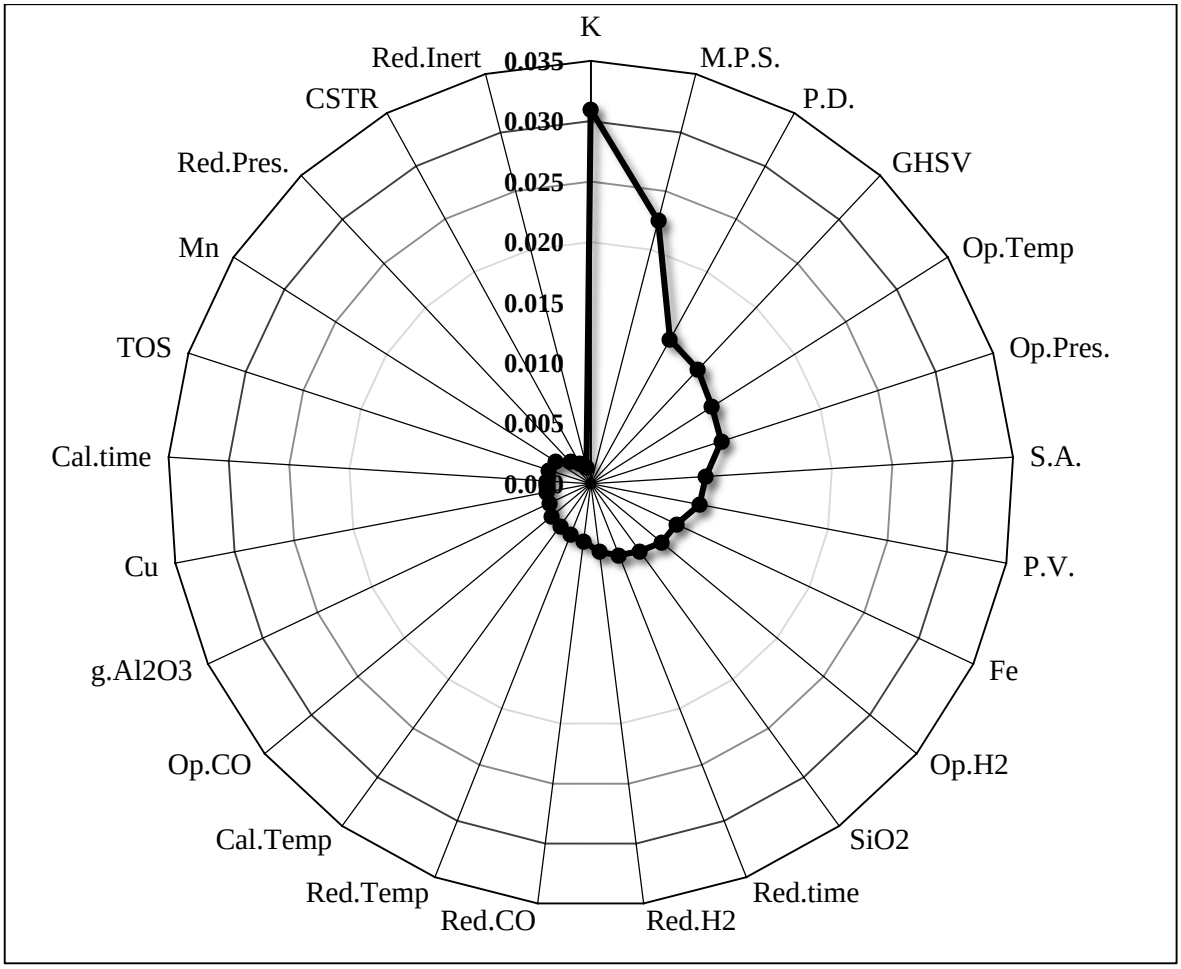

Figure 4.46. Relative importance of the attributes according to the Fe-based training set.

Figure 4.46 shows the relative importance of the attributes used in the optimum model. 25 out of 51 attributes were demonstrated to keep the figure interpretable. The attributes in the first quartile were observed to be significantly more effective on CO conversion than the remaining attributes. The most influential attribute was determined to be potassium (K) by far, which was an intriguing outcome since it was not even one of the top 25 important attributes in the model constructed by ANN, shown in Figure 4.20.

The top three important attributes were catalyst design related attributes, which was in accordance with the model constructed by ANN. Although the order of importance had

changed in two different models, they both indicating that catalyst design variables are slightly more important than operational variables.

Standardized residuals were plotted in Figure 4.47 with respect to three of the most important attributes, namely K loading (a), metal particle size (b), and pore diameter (c). Figure 4.47b and Figure 4.47c shows the random distribution of the residuals around zero, but a slight pattern was observed in Figure 4.47a. The residuals tended to be greater than zero between 0 and 2 wt.% K, but they attained negative values around 3 and 4 wt.% K. Therefore, although K was an essential constituent of the model, its effect was not captured sufficiently.

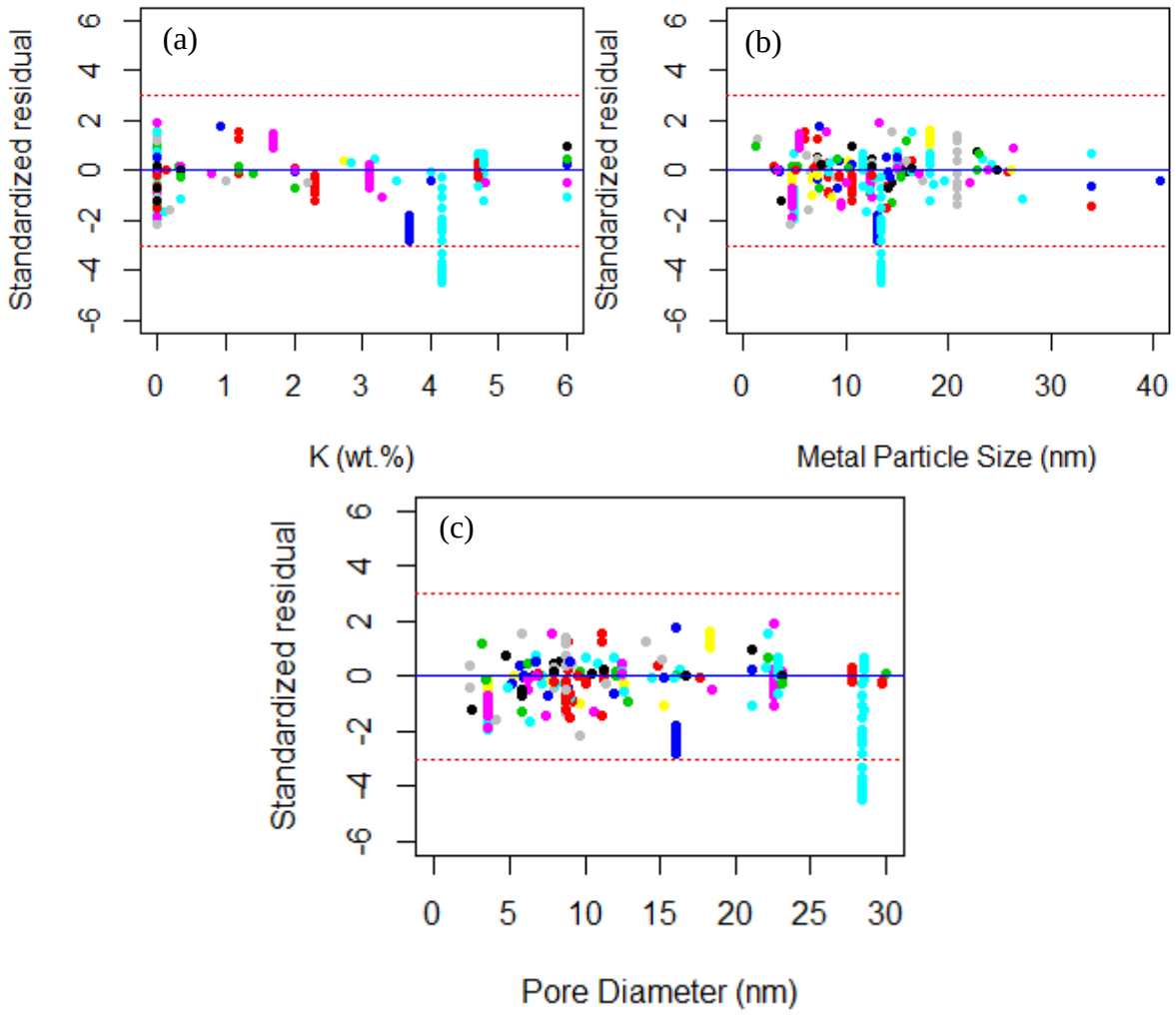

Figure 4.47. Standardized residuals vs. (a) K loading (b) metal particle size (c) pore diameter.

Figure 4.48 shows the standardized residuals with respect to time on stream. Blatant patterns were observed between the residuals with the same colors, indicating that the model failed to model the effect of time on stream sufficiently.

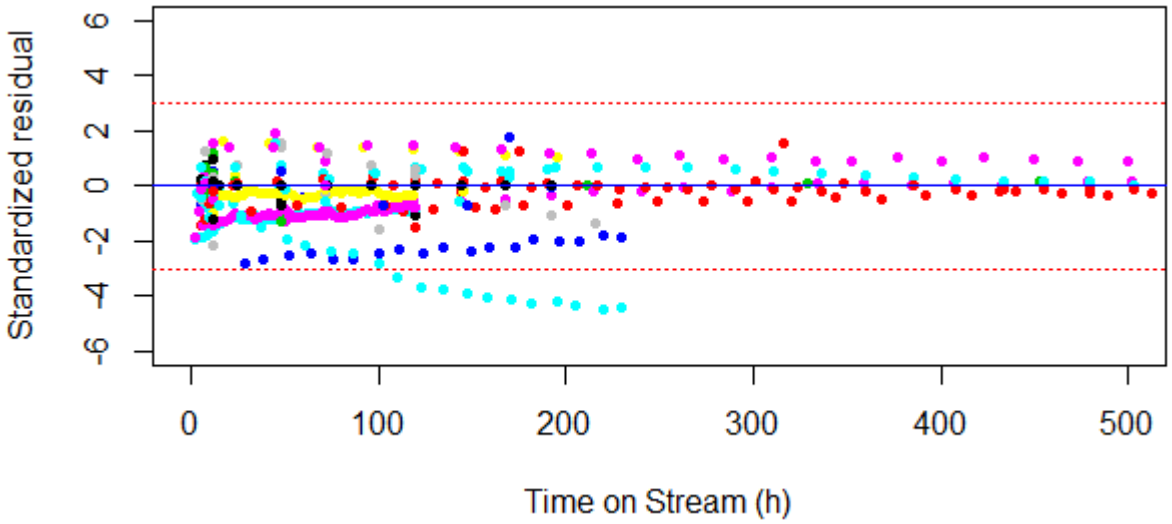

Figure 4.48. Standardized residuals vs. time on stream.

4.8.3. Low Temperature Subset

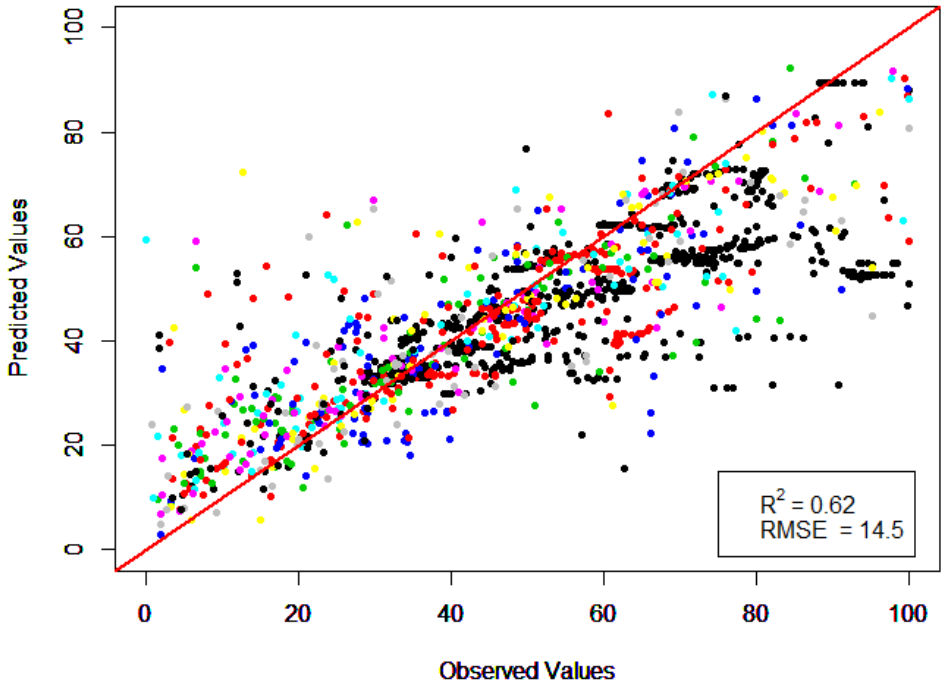

Figure 4.49. Predicted CO conversion vs. observed CO conversion for LT training set.

The dimensionality of the LT subset was reduced to 63 by following the procedure described in Figure 3.13. The best fitting forest was constructed with 72 trees, each of which was trained with 21 randomly selected attributes.

Figure 4.49 shows 10-fold cross validation predictions for the LT training set. The predictions of the random forest model was considerably more accurate than the ANN model. Yet, when it was compared with the random forest model constructed with the complete dataset, no significant improvement was observed. Although the R^2 value had increased, RMSE value had also increased slightly.

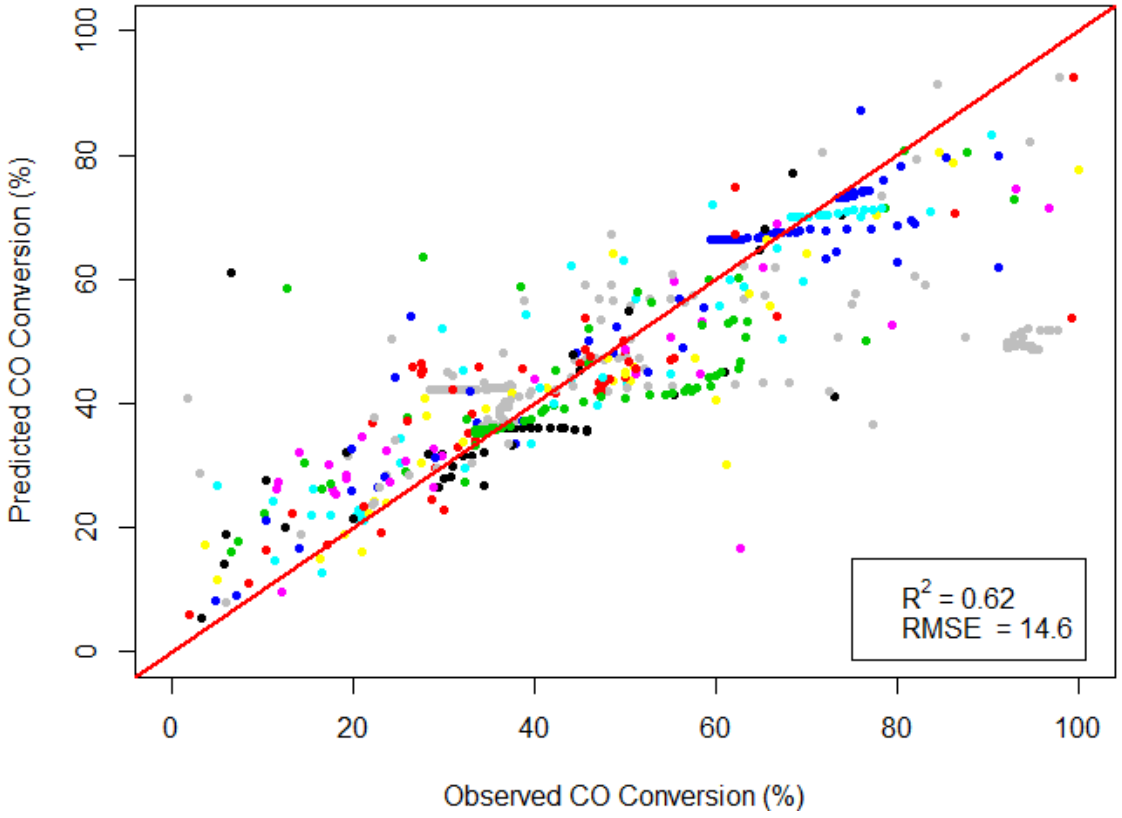

Figure 4.50. Predicted CO conversion vs. observed CO conversion for LT testing set.

Figure 4.50 shows the predictions of the testing set by the optimum random forest model. Both R^2 and RMSE values had improved, indicating that constructing an operating temperature-specific random forest model enhances LTFT predictions.

25 of the most important attributes constituting the LT random forest model were depicted in Figure 4.51. GHSV and operating temperature were determined as important

attributes related with the operating conditions. Both of the calcination parameters were designated to be crucial for the model, which was in consistence with the result of the ANN model. Metal particle size also appeared as one of the most important attributes. This was a reasonable outcome since in the literature, LTFT is usually conducted with Co or Ru based catalysts, both of which are active in their metallic phases, unlike Fe based catalysts.

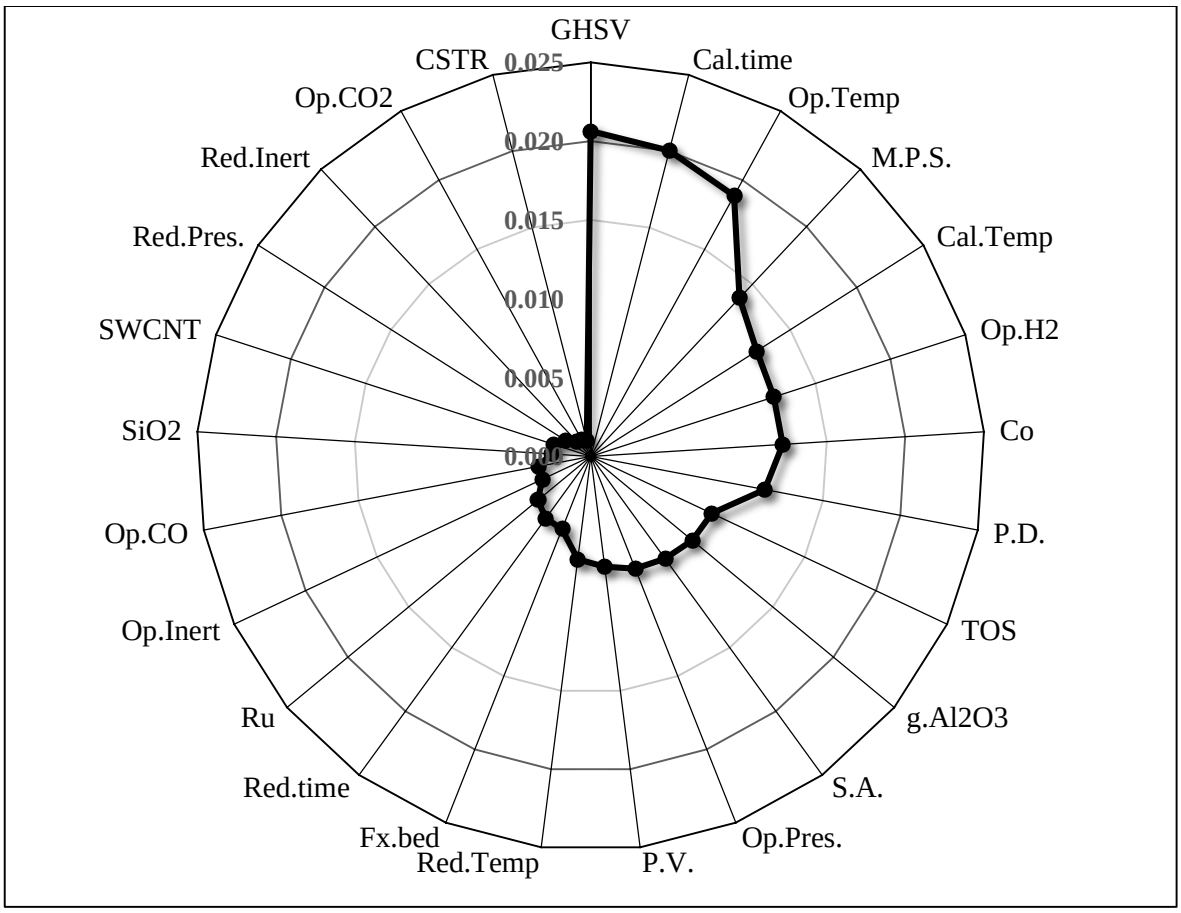

Figure 4.51. Relative importance of the attributes according to LT training set.

Figure 4.52 shows the standardized residuals of the LT random forest model with respect to GHSV (a), calcination time (b), and operating temperature (c). The residuals were randomly distributed around zero for all three attributes, indicating that the model could capture the effect of these attributes well.

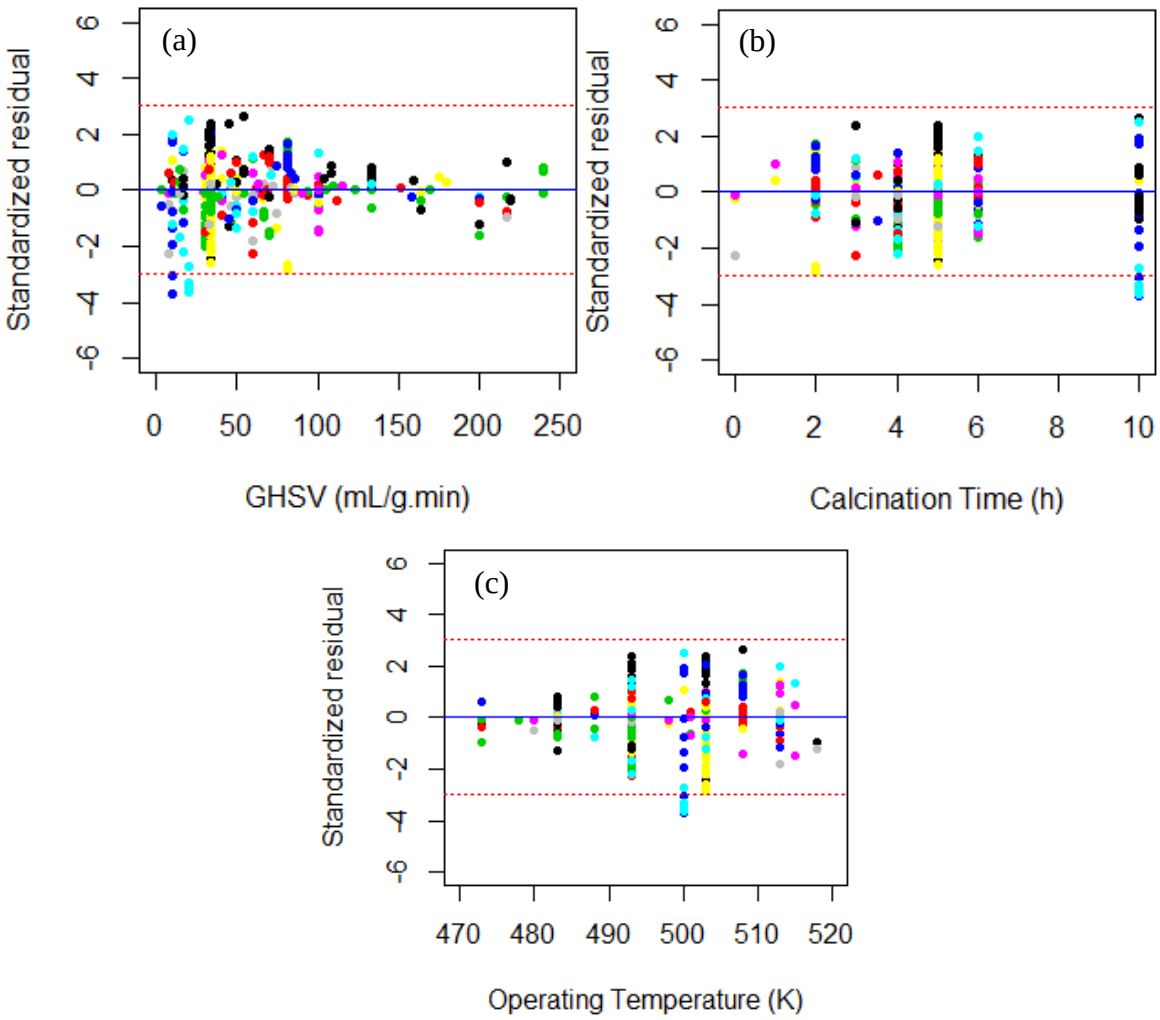

Figure 4.52. Standardized residuals vs. (a) GHSV (b) calcination time(c) operating temperature.

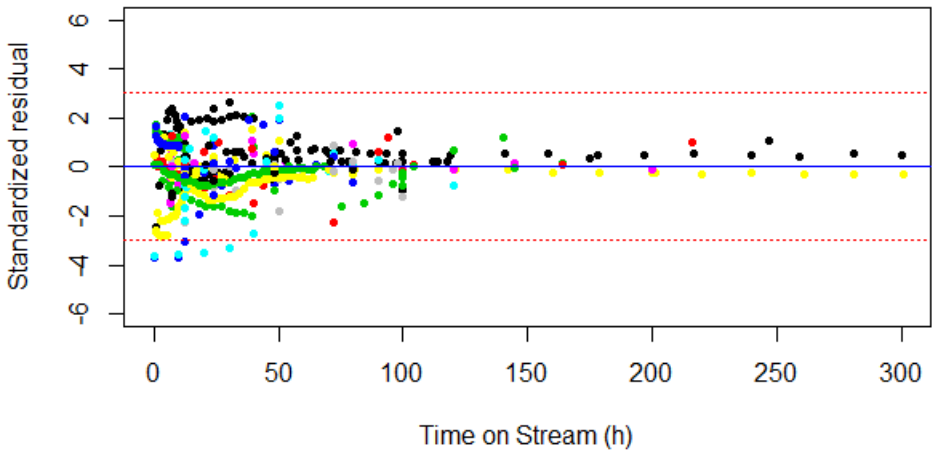

Figure 4.53. Standardized residuals time on stream.

Figure 4.53 shows the standardized residuals with respect to the time on stream. The residuals with the same colors showed distinct patterns, indicating that the model failed to model the behavior of time on stream.

4.8.4. High Temperature Subset

The dimensionality of the HT subset was reduced to 46 attributes after following the procedure described in Figure 3.13. The optimum forest was constructed with 42 trees, all of which were trained with 15 randomly selected attributes.

Figure 4.54 shows 10-fold cross validation predictions for the HT training set. The accuracy of the model was significantly higher than both its ANN counterpart and the random forest model trained by the complete dataset.

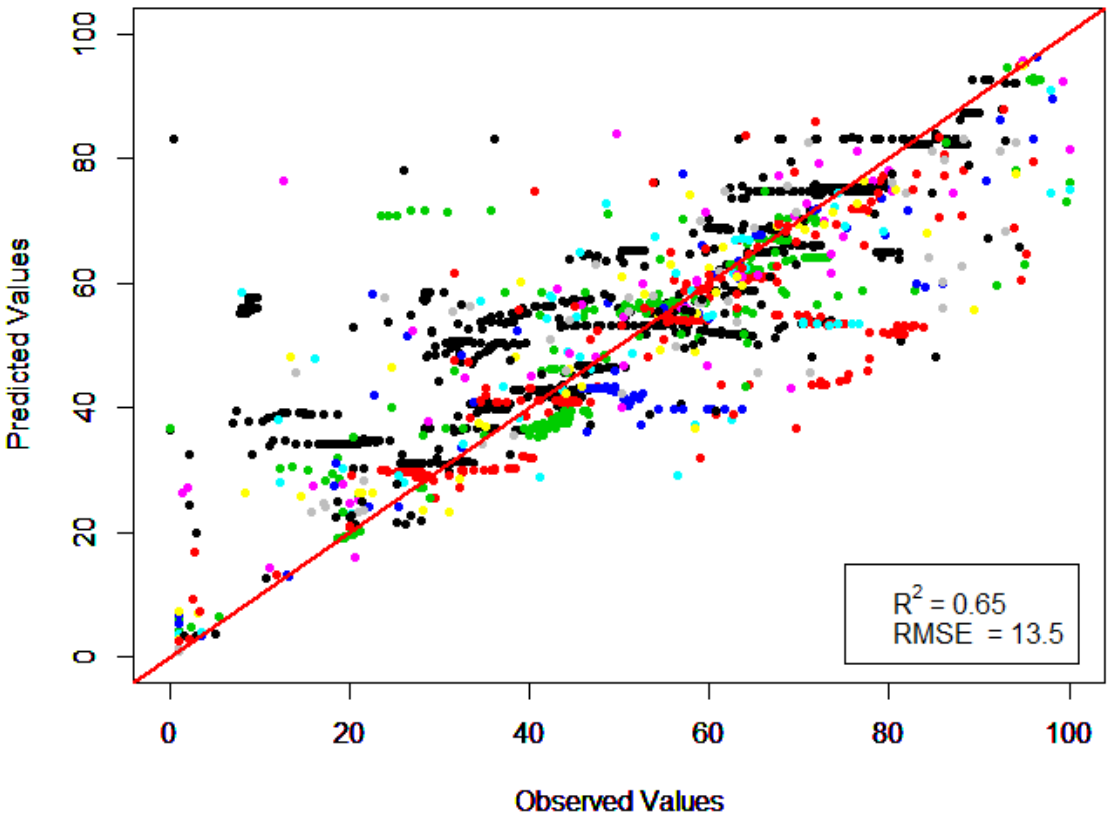

Figure 4.54. Predicted CO conversion vs. observed CO conversion for HT training set.

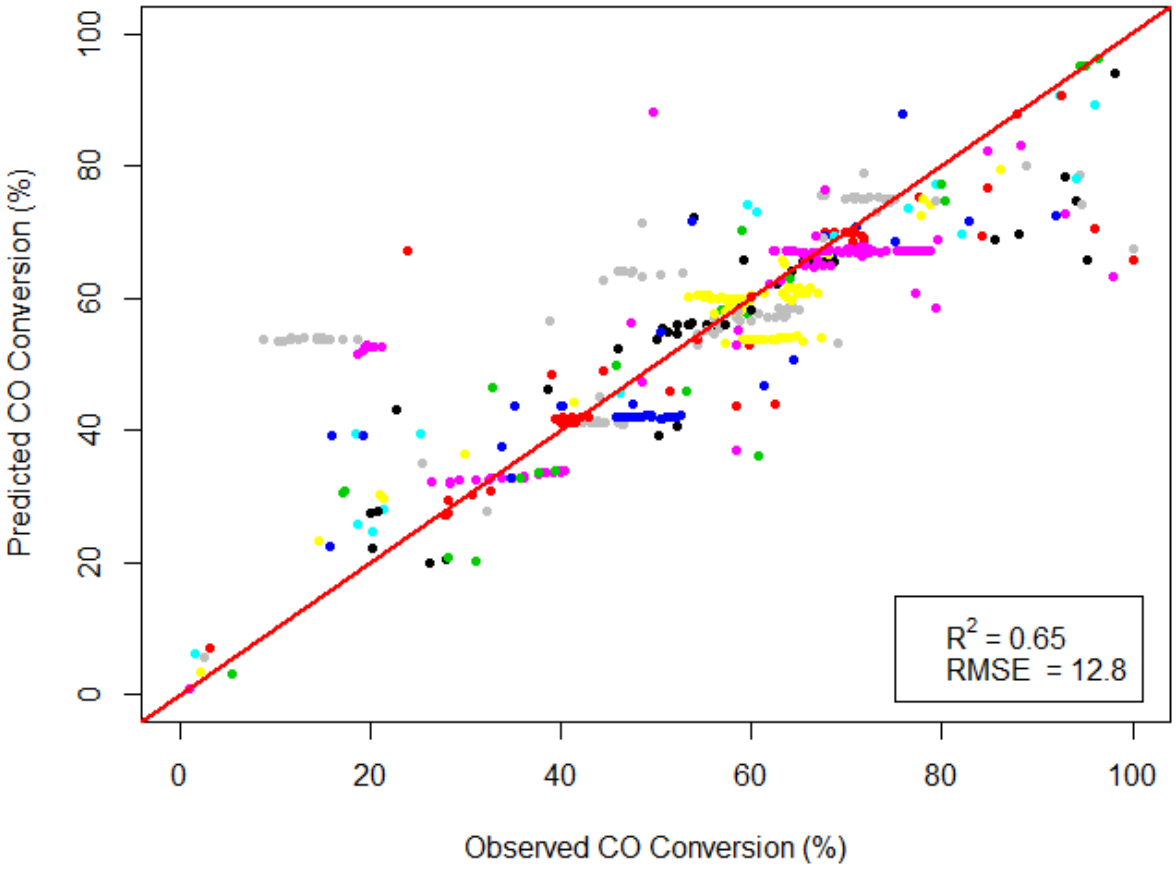

Figure 4.55. Predicted CO conversion vs. observed CO conversion for HT testing set.

Figure 4.55 shows the testing set predictions for the HT random forest model. The accuracy of the model had improved considerably, regarding the increased R^2 and decreased RMSE values. Therefore, it was concluded that splitting the data according to the operating temperature enhanced the predictions for HTFT conditions. Considering that LT based model also yielded improved results, the combined predictive power of the two models surpassed the performance of the model constructed with the complete dataset.

Figure 4.56 shows the relative importance attributes used in the HT random forest model. Operating temperature was designated as the most crucial attribute for the HTFT conditions. This result was unexpected since operating pressure is usually adjusted to tune the selectivity of the products, rather than maximizing CO conversion [92]. Limited number of study investigated the effect of pressure on CO conversion at HTFT conditions. Mirzaei *et al*. [94] and Feyzi *et al*. [42] reported that CO conversion decreases with increasing pressure above 350 °C. Also, PCA analysis in Figure 4.5b indicated that operating pressure

is mildly negatively correlated with CO conversion. Therefore, having operating pressure as a key attribute in the model was reasonable.

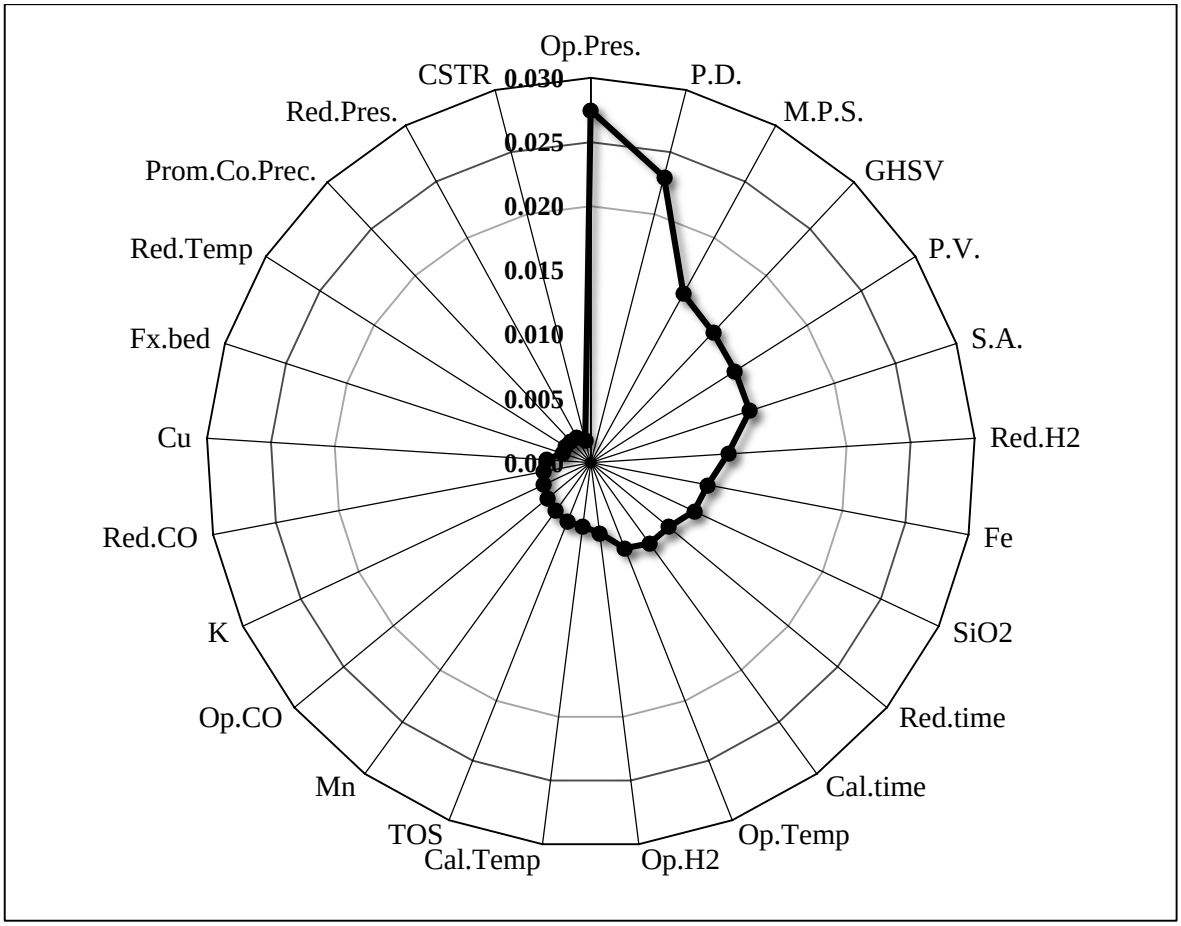

Figure 4.56. Relative importance of the attributes according to HT training set.

The importance analysis also indicated that all of the physical properties in the database were essential for the random forest model. This finding coincided with the HT ANN model, which yielded that promoter and support selection plays a key role in HTFT. Therefore, both of the models agreed that catalyst design variables are more important compared to operational variables for HT subset.

Figure 4.57 shows the standardized residuals with respect to operating pressure (a), pore diameter (b), and metal particle size (c). The residuals were observed to be randomly distributed around zero, indication of a good fit for these important attributes.

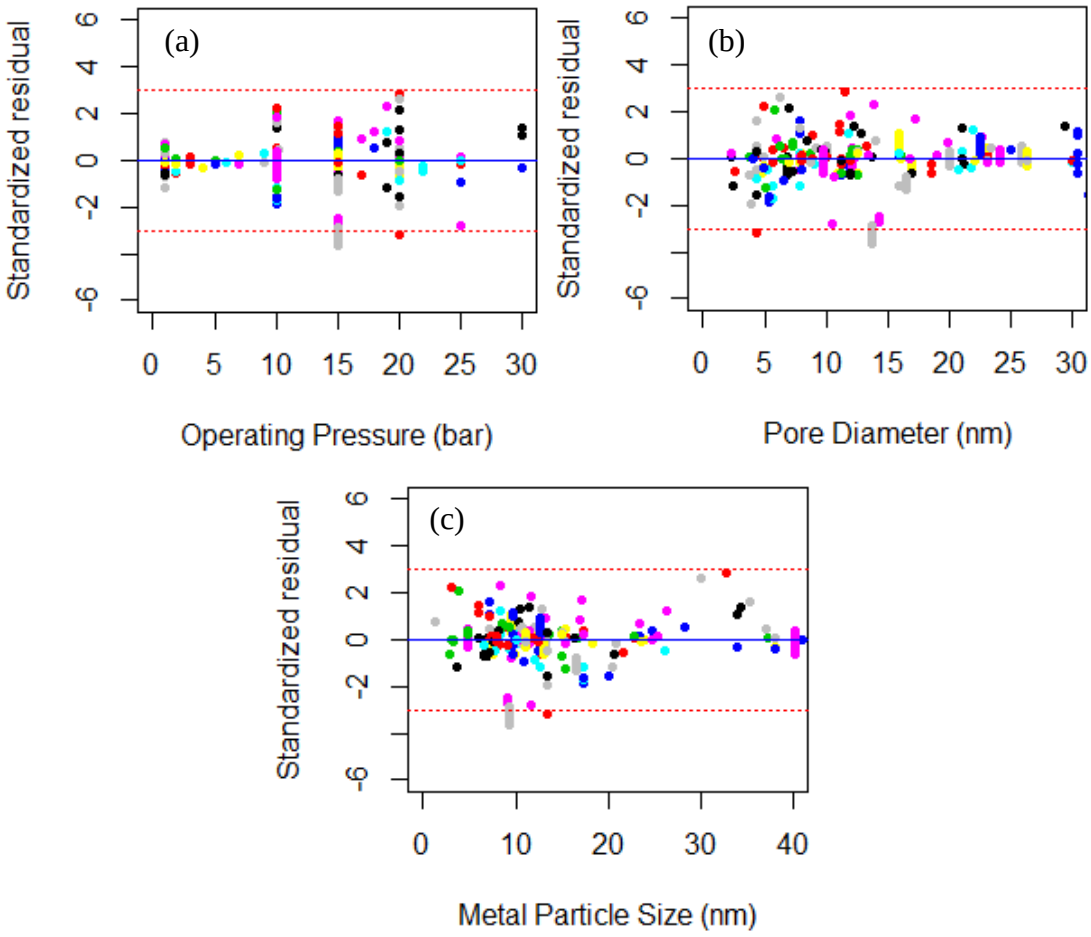

Figure 4.57. Standardized residuals vs. (a) operating pressure (b) pore diameter (c) metal particle size.

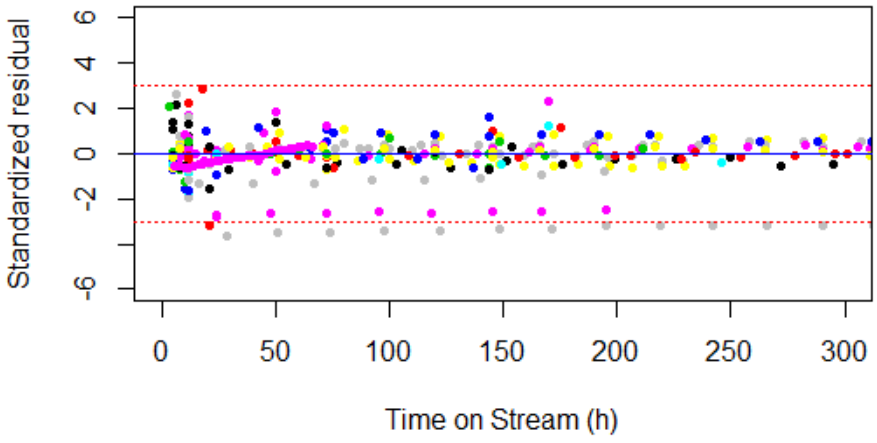

Figure 4.58. Standardized residuals vs. time on stream.

Figure 4.58 shows the standardized residuals with respect to time on stream. There were clear patterns between the residuals of the same experimental runs, indicating that this model failed to capture the effect of time on stream as well.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

In this study, FTS performance was investigated by collecting data from the published articles between 2005 and 2014 from the databases of ScienceDirect, Wiley Online Library, and ACS Publications. 4098 data points were extracted from 171 research papers, and all the instances were represented in a 72 dimensional space. PCA, multiple linear regression, ANN, decision tree, and random forest algorithms were used to analyze the constructed database to reveal hidden relationships between catalyst design variables, physical properties, operating conditions, and CO conversion. All the computational work was carried out in R 3.1.2 environment.

Since the physical properties in the database had missing values, the procedure started with completing them by benefiting from their interrelationship and their relationship with the catalyst design variables. ANN algorithm was used to construct models based on the known instances to predict the missing values in the physical properties. An iterative procedure was employed to model surface area, pore volume, pore diameter, and metal particle size. Each physical property was modelled separately in each iteration, and the preceding properties were included in the input variables prior to modelling. The R^2 values for the randomly selected testing sets were 0.32, 0.55, 0.63, and 0.76 for surface area, pore volume, pore diameter, and metal particle size, respectively. The missing values were completed by using the model trained by the known instances.

PCA algorithm was used to analyze the correlations between CO conversion and the physical properties as well as the operating variables on two dimensional plots. It was observed that operating temperature and pore diameter of the catalyst are highly correlated with CO conversion. On the other hand, the correlation between the CO conversion and GHSV, TOS, and metal particle size were mildly negatively correlated.

Multiple linear regression model was constructed to predict the numerical value of CO conversion. The model was unsuccessful; therefore, no further effort was spent to improve

the model or to interpret the results, and more complex machine learning algorithms were constructed for further analyses.

An ANN model was constructed with 15 hidden units in a single layer, all of which were activated by logistic sigmoid function and trained by standard backpropagation algorithm. The R^2 value of the testing set was computed as 0.40, which indicated that the accuracy of the model was unsatisfactory. The importance analysis on the resultant model revealed that operating temperature, pore volume, and hydrogen composition in the feed gas are the most important attributes in FTS process. Sensitivity analysis indicated that an increase in those attributes mostly increases the CO conversion. The residual analysis showed that although the model could capture the behavior of these significant attributes, it failed to represent the effect of time on stream on CO conversion.

Since the predictive power of the model was weak; Co-based, Fe-based, LT and HT subsets were used to construct sub-models. All of these models yielded improved accuracies with R^2 values of 0.50, 0.53, 0.50, and 0.43, respectively on the testing sets. The importance analysis revealed that catalyst design variables were more dominant in the Fe-based subset, whereas operating conditions were more important in the Co-based subset.

A decision tree was trained by using the complete dataset to bring out meaningful and interpretable relations between the input variables and CO conversion. However, the prediction accuracy of the model was poor with a misclassification error of 40.4% on the testing set. Therefore, any deductions based on such a weak model would be erroneous and misleading.

Random forest was the last algorithm used for modelling the CO conversion. The R^2 value of the testing set was 0.60, which indicated that its predictive power is significantly higher than its ANN counterpart. The importance analysis on the random forest model validated the outcomes of the ANN model by marking similar attributes as important.

The R^2 values for the Co-based, Fe-based, LT, and HT random forest models on the testing sets were 0.66, 0.64, 0.62, and 0.65, respectively. The improved accuracy values of these models compared to the ANN models indicated that random forest algorithm is much

more suitable for modelling CO conversion in FTS. Also, it should be noted that the importance analyses of these models yielded consistent results with the ANN models.

5.2. Recommendations

This study can be further improved by extending the scope to investigate the optimum FTS activity under specified conditions. Both random forest and ANN models can be integrated with an optimization algorithm, such as genetic algorithm, to search for the optimum values of the attributes that maximize CO conversion while some of the attributes are specified. For example, if a researcher wanted to investigate the conditions that maximizes CO conversion for a catalyst he had developed, he could benefit from that model to narrow down the possibilities for the operating conditions.

Another improvement might be including the selectivity values for CO_2, CH_4, and hydrocarbon chains as additional output variables. Since most of the researchers try improving the yield for longer hydrocarbon chains, predicting and optimizing the yields might be more beneficial.

REFERENCES

1. British Petroleum, *BP Statistical Review of World Energy*, 64th edition, B.P., London, 2015.

2. Owen, N. A., O. R. Inderwildi, and D. A. King, "The Status of Conventional World Oil Reserves-Hype or Cause for Concern?", *Energy Policy*, Vol. 38, No. 8, pp. 4743-4749, 2010.

3. Zhang, Q., W. Deng, and Y. Wang, "Recent Advances in Understanding the Key Catalyst Factors for Fischer-Tropsch Synthesis", *Journal of Energy Chemistry*, Vol. 22, No. 1, pp. 27-38, 2013.

4. Galvis, H. M. T. and K. P. D. Jong, "Catalysts for Production of Lower Olefins from Synthesis Gas : A Review", *ACS Catalysis*, Vol. 3, pp. 2130-2149, 2013.

5. Lu, Y. and T. Lee, "Influence of the Feed Gas Composition on the Fischer-Tropsch Synthesis in Commercial Operations", *Journal of Natural Gas Chemistry*, Vol. 16, No. 4, pp. 329-341, 2007.

6. Zhang, Q., J. Kang, and Y. Wang, "Development of Novel Catalysts for Fischer-Tropsch Synthesis: Tuning the Product Selectivity", *ChemCatChem*, Vol. 2, No. 9, pp. 1030-1058, 2010.

7. Alpaydın, E., *Introduction to Machine Learning*, 2nd ed, Cambridge, Massachusetts, The MIT Press, 2010.

8. Günay, M. E. and R. Yildirim, "Modeling Preferential Co Oxidation over Promoted Au/Al_2O_3 Catalysts Using Decision Trees and Modular Neural Networks", *Chemical Engineering Research and Design*, Vol. 91, No. 5, pp. 874-882, 2013.

9. Günay, M. E. and R. Yildirim, "Developing Global Reaction Rate Model for Co Oxidation over Au Catalysts from Past Data in Literature Using Artificial Neural Networks", *Applied Catalysis A: General*, Vol. 468, pp. 395-402, 2013.

10. Shimura, K., T. Miyazawa, T. Hanaoka, and S. Hirata, "Preparation of Co/Al_2O_3 Catalyst for Fischer-Tropsch Synthesis: Combination of Impregnation Method and Homogeneous Precipitation Method", *Applied Catalysis A: General*, Vol. 475, pp. 1-9, 2014.

11. Trépanier, M., A. K. Dalai, and N. Abatzoglou, "Synthesis of CNT-Supported Cobalt Nanoparticle Catalysts Using a Microemulsion Technique: Role of Nanoparticle Size on Reducibility, Activity and Selectivity in Fischer-Tropsch Reactions", *Applied Catalysis A: General*, Vol. 374, No. 1-2, pp. 79-86, 2010.

12. Dalai, A. K. and B. H. Davis, "Fischer-Tropsch Synthesis: A Review of Water Effects on the Performances of Unsupported and Supported Co Catalysts", *Applied Catalysis A: General*, Vol. 348, No. 1, pp. 1-15, 2008.

13. Lögdberg, S., D. Tristantini, Ø. Borg, L. Ilver, B. Gevert, S. Järås, E. A. Blekkan, and A. Holmen, "Hydrocarbon Production Via Fischer-Tropsch Synthesis from H_2-Poor Syngas over Different Fe-Co/γ-Al_2O_3 Bimetallic Catalysts", *Applied Catalysis B: Environmental*, Vol. 89, No. 1-2, pp. 167-182, 2009.

14. Dry, M. E., "High Quality Diesel Via the Fischer-Tropsch Process - A Review", *Journal of Chemical Technology and Biotechnology*, Vol. 77, No. 1, pp. 43-50, 2002.

15. Trépanier, M., A. Tavasoli, A. K. Dalai, and N. Abatzoglou, "Co, Ru and K Loadings Effects on the Activity and Selectivity of Carbon Nanotubes Supported Cobalt Catalyst in Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 353, No. 2, pp. 193-202, 2009.

16. Okabe, K., M. Wei, and H. Arakawa, "Fischer-Tropsch Synthesis over Cobalt Catalysts Supported on Mesoporous Metallo-Silicates", Vol. 58, No. 12, pp. 822-828, 2003.

17. Khodaei, M. M., M. Feyzi, J. Shahmoradi, and M. Joshaghani, "The Sol-Gel Derived Co-Mn/TiO_2 Catalysts for Light Olefins Production", *Journal of Fuel Chemistry and Technology*, Vol. 42, No. 2, pp. 212-218, 2014.

18. Tavasoli, A., R. M. M. Abbaslou, M. Trepanier, and A. K. Dalai, "Fischer-Tropsch Synthesis over Cobalt Catalyst Supported on Carbon Nanotubes in a Slurry Reactor", *Applied Catalysis A: General*, Vol. 345, No. 2, pp. 134-142, 2008.

19. Bae, J. W., S. M. Kim, S. H. Kang, K. V. R. Chary, Y. J. Lee, H. J. Kim, and K. W. Jun, "Effect of Support and Cobalt Precursors on the Activity of Co/$AlPO_4$ Catalysts in Fischer-Tropsch Synthesis", *Journal of Molecular Catalysis A: Chemical*, Vol. 311, No. 1-2, pp. 7-16, 2009.

20. Yu, L., X. Liu, Y. Fang, C. Wang, and Y. Sun, "Highly Active Co/SiC Catalysts with Controllable Dispersion and Reducibility for Fischer-Tropsch Synthesis", *Fuel*, Vol. 112, No. pp. 483-488, 2013.

21. Kang, S. H., H. M. Koo, a. R. Kim, D. H. Lee, J. H. Ryu, Y. D. Yoo, and J. W. Bae, "Correlation of the Amount of Carbonaceous Species with Catalytic Performance on Iron-Based Fischer-Tropsch Catalysts", *Fuel Processing Technology*, Vol. 109, No. pp. 141-149, 2013.

22. Riedel, T., H. Schulz, G. Schaub, K. W. Jun, J. S. Hwang, and K. W. Lee, "Fischer-Tropsch on Iron with H_2/CO and H_2/CO_2 as Synthesis Gases: The Episodes of Formation of the Fischer-Tropsch Regime and Construction of the Catalyst", *Topics in Catalysis*, Vol. 26, No. 1-4, pp. 41-54, 2003.

23. Torres Galvis, H. M., A. C. J. Koeken, J. H. Bitter, T. Davidian, M. Ruitenbeek, A. I. Dugulan, and K. P. De Jong, "Effects of Sodium and Sulfur on Catalytic Performance of Supported Iron Catalysts for the Fischer-Tropsch Synthesis of Lower Olefins", *Journal of Catalysis*, Vol. 303, pp. 22-30, 2013.

24. Baranak, M., B. Gürünlü, A. Sarioğlan, Ö. Ataç, and H. Atakül, "Low Acidity ZSM-5 Supported Iron Catalysts for Fischer-Tropsch Synthesis", *Catalysis Today*, Vol. 207, pp. 57-64, 2013.

25. Nurunnabi, M., K. Murata, K. Okabe, M. Inaba, and I. Takahara, "Performance and Characterization of Ru/Al_2O_3 and Ru/SiO_2 Catalysts Modified with Mn for Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 340, No. 2, pp. 203-211, 2008.

26. Griboval-Constant, A., A. Butel, V. V. Ordomsky, P. A. Chernavskii, and A. Y. Khodakov, "Cobalt and Iron Species in Alumina Supported Bimetallic Catalysts for Fischer-Tropsch Reaction", *Applied Catalysis A: General*, Vol. 481, pp. 116-126, 2014.

27. Song, S. H., S. B. Lee, J. W. Bae, P. S. Sai Prasad, and K. W. Jun, "Influence of Ru Segregation on the Activity of Ru-Co/γ-Al_2O_3 During FT Synthesis: A Comparison with That of Ru-Co/SiO_2 Catalysts", *Catalysis Communications*, Vol. 9, No. 13, pp. 2282-2286, 2008.

28. Ngamcharussrivichai, C., X. Liu, X. Li, T. Vitidsant, and K. Fujimoto, "An Active and Selective Production of Gasoline-Range Hydrocarbons over Bifunctional Co-Based Catalysts", *Fuel*, Vol. 86, No. 1-2, pp. 50-59, 2007.

29. Mirzaei, A. A., S. Vahid, and H. O. Torshizi, "Effect of Support and Promoter on the Catalytic Performance and Structural Properties of the Fe-Co-Ni Catalysts for Co Hydrogenation", *Journal of Natural Gas Science and Engineering*, Vol. 15, No. 4, pp. 106-117, 2013.

30. Rane, S., Ø. Borg, J. Yang, E. Rytter, and A. Holmen, "Effect of Alumina Phases on Hydrocarbon Selectivity in Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 388, No. 1-2, pp. 160-167, 2010.

31. Rane, S., O. Borg, E. Rytter, and A. Holmen, "Relation Between Hydrocarbon Selectivity and Cobalt Particle Size for Alumina Supported Cobalt Fischer-Tropsch Catalysts", *Applied Catalysis A: General*, Vol. 437-438, pp. 10-17, 2012.

32. Zhang, X., Y. Liu, G. Liu, K. Tao, Q. Jin, F. Meng, D. Wang, and N. Tsubaki, "Product Distributions Including Hydrocarbon and Oxygenates of Fischer-Tropsch Synthesis over Mesoporous MnO_2-Supported Fe Catalyst", *Fuel*, Vol. 92, No. 1, pp. 122-129, 2012.

33. Ding, M., M. Qiu, J. Liu, Y. Li, T. Wang, L. Ma, and C. Wu, "Influence of Manganese Promoter on Co-Precipitated Fe-Cu Based Catalysts for Higher Alcohols Synthesis", *Fuel*, Vol. 109, pp. 21-27, 2013.

34. Pan, Z. and D. B. Bukur, "Fischer-Tropsch Synthesis on Co/ZnO Catalyst - Effect of Pretreatment Procedure", *Applied Catalysis A: General*, Vol. 404, No. 1-2, pp. 74-80, 2011.

35. Pan, Z., M. Parvari, and D. B. Bukur, "Fischer-Tropsch Synthesis on Co/ZnO - Two Step Activation Procedure for Improved Performance", *Applied Catalysis A: General*, Vol. 480, pp. 79-85, 2014.

36. Zhang, X., H. Su, and X. Yang, "Catalytic Performance of a Three-Dimensionally Ordered Macroporous Co/ZrO_2 Catalyst in Fischer–Tropsch Synthesis", *Journal of Molecular Catalysis A: Chemical*, Vol. 360, pp. 16-25, 2012.

37. Sun, Y., Q. W. Sun, F. K. Jiang, J. S. Liu, and Z. S. Zhang, "Effects of Calcination and Reduction Temperatures on the Performance of $Co\text{-}Pt\text{-}ZrO_2/\gamma\text{-}Al_2O_3$ Catalysts for Fischer-Tropsch Synthesis", *Journal of Fuel Chemistry and Technology*, Vol. 40, No. 1, pp. 54-58, 2012.

38. Qing, M., Y. Yang, B. Wu, J. Xu, C. Zhang, P. Gao, and Y. Li, "Modification of Fe-SiO_2 Interaction with Zirconia for Iron-Based Fischer-Tropsch Catalysts", *Journal of Catalysis*, Vol. 279, No. 1, pp. 111-122, 2011.

39. Park, S. J., J. W. Bae, J. H. Oh, K. V. R. Chary, P. S. S. Prasad, K. W. Jun, and Y. W. Rhee, "Influence of Bimodal Pore Size Distribution of $Ru/Co/ZrO_2\text{-}Al_2O_3$ During Fischer-Tropsch Synthesis in Fixed-Bed and Slurry Reactor", *Journal of Molecular Catalysis A: Chemical*, Vol. 298, No. 1-2, pp. 81-87, 2009.

40. Gnanamani, M. K., M. C. Ribeiro, W. Ma, W. D. Shafer, G. Jacobs, U. M. Graham, and B. H. Davis, "Fischer-Tropsch Synthesis: Metal-Support Interfacial Contact Governs Oxygenates Selectivity over CeO_2 Supported Pt-Co Catalysts", *Applied Catalysis A: General*, Vol. 393, No. 1-2, pp. 17-23, 2011.

41. Dai, X., "Characterization and Catalytic Performance of CeO_2-Co/SiO_2 Catalyst for Fischer-Tropsch Synthesis Using Nitrogen-Diluted Synthesis Gas over a Laboratory Scale Fixed-Bed Reactor", Journal of Natural Gas Chemistry, Vol. 17, pp. 17-23, 2008.

42. Feyzi, M., A. A. Mirzaei, and H. R. Bozorgzadeh, "Effects of Preparation and Operation Conditions on Precipitated Iron Nickel Catalysts for Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 19, No. 3, pp. 341-353, 2010.

43. Cai, Z., J. Li, K. Liew, and J. Hu, "Effect of La_2O_3-Dopping on the Al_2O_3 Supported Cobalt Catalyst for Fischer-Tropsch Synthesis", *Journal of Molecular Catalysis A: Chemical*, Vol. 330, No. 1-2, pp. 10-17, 2010.

44. Sartipi, S., J. E. Van Dijk, J. Gascon, and F. Kapteijn, "Toward Bifunctional Catalysts for the Direct Conversion of Syngas to Gasoline Range Hydrocarbons: H-ZSM-5 Coated Co Versus H-ZSM-5 Supported Co", *Applied Catalysis A: General*, Vol. 456, pp. 11-22, 2013.

45. Xiong, H., L. L. Jewell, and N. J. Coville, "Shaped Carbons as Supports for the Catalytic Conversion of Syngas to Clean Fuels", *ACS Catalysis*, Vol. 5, pp. 2640-2658, 2015.

46. Liu, Y., J. Luo, M. Girleanu, O. Ersen, C. Pham-Huu, and C. Meny, "Efficient Hierarchically Structured Composites Containing Cobalt Catalyst for Clean Synthetic Fuel Production from Fischer–Tropsch Synthesis", *Journal of Catalysis*, Vol. 318, pp. 179-192, 2014.

47. Liu, Y., T. Dintzer, O. Ersen, and C. Pham-Huu, "Carbon Nanotubes Decorated A-Al_2O_3 Containing Cobalt Nanoparticles for Fischer-Tropsch Reaction", *Journal of Energy Chemistry*, Vol. 22, No. 2, pp. 279-289, 2013.

48. Fu, T. and Z. Li, "Highly Dispersed Cobalt on N-Doped Carbon Nanotubes with Improved Fischer-Tropsch Synthesis Activity", *Catalysis Communications*, Vol. 47, No. 3, pp. 54-57, 2014.

49. Fu, T., R. Liu, J. Lv, and Z. Li, "Influence of Acid Treatment on N-Doped Multi-Walled Carbon Nanotube Supports for Fischer-Tropsch Performance on Cobalt Catalyst", *Fuel Processing Technology*, Vol. 122, pp. 49-57, 2014.

50. Xiong, H., M. a. M. Motchelaho, M. Moyo, L. L. Jewell, and N. J. Coville, "Cobalt Catalysts Supported on a Micro-Coil Carbon in Fischer-Tropsch Synthesis: A Comparison with CNTs and CNFs", *Catalysis Today*, Vol. 214, pp. 50-60, 2013.

51. Xiong, K., Y. Zhang, J. Li, and K. Liew, "Catalytic Properties of Ru Nanoparticles Embedded on Ordered Mesoporous Carbon with Different Pore Size in Fischer-Tropsch Synthesis", *Journal of Energy Chemistry*, Vol. 22, No. 4, pp. 560-566, 2013.

52. Honsho, T. O., T. Kitano, T. Miyake, and T. Suzuki, "Fischer-Tropsch Synthesis over Co-Loaded Oxidized Diamond Catalyst", *Fuel*, Vol. 94, pp. 170-177, 2012.

53. Xiong, H., M. Moyo, M. K. Rayner, L. L. Jewell, D. G. Billing, and N. J. Coville, "Autoreduction and Catalytic Performance of a Cobalt Fischer-Tropsch Synthesis Catalyst Supported on Nitrogen-Doped Carbon Spheres", *ChemCatChem*, Vol. 2, No. 5, pp. 514-518, 2010.

54. Li, T., H. Wang, Y. Yang, H. Xiang, and Y. Li, "Study on an Iron-Nickel Bimetallic Fischer-Tropsch Synthesis Catalyst", *Fuel Processing Technology*, Vol. 118, pp. 1-8, 2014.

55. Feyzi, M., M. M. Khodaei, and J. Shahmoradi, "Effect of Preparation and Operation Conditions on the Catalytic Performance of Cobalt-Based Catalysts for Light Olefins Production", *Fuel Processing Technology*, Vol. 93, No. 1, pp. 90-98, 2012.

56. Li, J., C. Zhang, X. Cheng, M. Qing, J. Xu, B. Wu, Y. Yang, and Y. Li, "Effects of Alkaline-Earth Metals on the Structure, Adsorption and Catalytic Behavior of Iron-

Based Fischer-Tropsch Synthesis Catalysts", *Applied Catalysis A: General*, Vol. 464-465, pp. 10-19, 2013.

57. Bromfield, T. C. and N. J. Coville, "The Effect of Sulfide Ions on a Precipitated Iron Fischer–Tropsch Catalyst", *Applied Catalysis A: General*, Vol. 186, No. 1-2, pp. 297-307, 1999.

58. Torres-Galvis, H. M., A. C. J. Koeken, J. H. Bitter, T. Davidian, M. Ruitenbeek, A. I. Dugulan, and K. P. De Jong, "Effect of Precursor on the Catalytic Performance of Supported Iron Catalysts for the Fischer-Tropsch Synthesis of Lower Olefins", *Catalysis Today*, Vol. 215, pp. 95-102, 2013.

59. Bao, Z., K. Xiao, X. Qi, X. Wang, L. Zhong, K. Fang, M. Lin, and Y. Sun, "Higher Alcohol Synthesis over Cu-Fe Composite Oxides with High Selectivity to C_2+OH", *Journal of Energy Chemistry*, Vol. 22, No. 1, pp. 107-113, 2013.

60. Lohitharn, N., J. G. Goodwin, and E. Lotero, "Fe-Based Fischer-Tropsch Synthesis Catalysts Containing Carbide-Forming Transition Metal Promoters", *Journal of Catalysis*, Vol. 255, No. 1, pp. 104-113, 2008.

61. Lohitharn, N. and J. G. Goodwin, "Impact of Cr, Mn and Zr Addition on Fe Fischer-Tropsch Synthesis Catalysis: Investigation at the Active Site Level Using SSITKA", *Journal of Catalysis*, Vol. 257, No. 1, pp. 142-151, 2008.

62. Miyazawa, T., T. Hanaoka, K. Shimura, and S. Hirata, "Mn and Zr Modified Co/SiO_2 Catalysts Development in Slurry-Phase Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 467, No. 3, pp. 47-54, 2013.

63. Wang, G., K. Zhang, P. Liu, H. Hui, and Y. Tan, "Synthesis of Light Olefins from Syngas over Fe-Mn-V-K Catalysts in the Slurry Phase", *Journal of Industrial and Engineering Chemistry*, Vol. 19, No. 3, pp. 961-965, 2013.

64. Zhang, J. L., L. H. Ma, S. B. Fan, T. S. Zhao, and Y. H. Sun, "Synthesis of Light Olefins from Co Hydrogenation over Fe-Mn Catalysts: Effect of Carburization Pretreatment", *Fuel*, Vol. 109, pp. 116-123, 2013.

65. Yang, Z., X. Pan, J. Wang, and X. Bao, "Fe Nanoparticles Confined inside CNT for Light Olefin Synthesis from Syngas: Effects of Mn and K Additives", *Catalysis Today*, Vol. 186, No. 1, pp. 121-127, 2012.

66. Qin, S., C. Zhang, J. Xu, Y. Yang, H. Xiang, and Y. Li, "Fe-Mo Interactions and Their Influence on Fischer-Tropsch Synthesis Performance", *Applied Catalysis A: General*, Vol. 392, No. 1-2, pp. 118-126, 2011.

67. Malek Abbaslou, R. M., J. Soltan, and A. K. Dalai, "Iron Catalyst Supported on Carbon Nanotubes for Fischer-Tropsch Synthesis: Effects of Mo Promotion", *Fuel*, Vol. 90, No. 3, pp. 1139-1144, 2011.

68. Qin, S., C. Zhang, J. Xu, B. Wu, H. Xiang, and Y. Li, "Effect of Mo Addition on Precipitated Fe Catalysts for Fischer-Tropsch Synthesis", *Journal of Molecular Catalysis A: Chemical*, Vol. 304, No. 1-2, pp. 128-134, 2009.

69. Chen, L., G. Song, Y. Fu, and J. Shen, "The Effects of Promoters of K and Zr on the Mesoporous Carbon Supported Cobalt Catalysts for Fischer-Tropsch Synthesis", *Journal of Colloid and Interface Science*, Vol. 368, No. 1, pp. 456-461, 2012.

70. Parnian, M. J., A. A. Khodadadi, A. Taheri Najafabadi, and Y. Mortazavi, "Preferential Chemical Vapor Deposition of Ruthenium on Cobalt with Highly Enhanced Activity and Selectivity for Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 470, pp. 221-231, 2014.

71. Akia, M., S. M. Alavi, and Z. F. Yan, "Dehyrogenation Catalysts of Higher Normal Paraffins on a Nanocrystalline Al_2O_3 Support: Different Impregnation Sequences", *Applied Surface Science*, Vol. 2, pp. 64-73, 2011.

72. Yao, M., N. Yao, Y. Shao, Q. Han, C. Ma, C. Yuan, C. Li, and X. Li, "New Insight into the Activity of ZSM-5 Supported Co and CoRu Bifunctional Fischer-Tropsch Synthesis Catalyst", *Chemical Engineering Journal*, Vol. 239, No. pp. 408-415, 2014.

73. Wang, S., Q. Yin, J. Guo, B. Ru, and L. Zhu, "Improved Fischer-Tropsch Synthesis for Gasoline over Ru, Ni Promoted Co/HZSM-5 Catalysts", *Fuel*, Vol. 108, pp. 597-603, 2013.

74. Pirola, C., M. Scavini, F. Galli, S. Vitali, A. Comazzi, F. Manenti, and P. Ghigna, "Fischer-Tropsch Synthesis: Exafs Study of Ru and Pt Bimetallic Co Based Catalysts", *Fuel*, Vol. 132, pp. 62-70, 2014.

75. Ishida, T., T. Yanagihara, X. Liu, H. Ohashi, A. Hamasaki, T. Honma, H. Oji, T. Yokoyama, and M. Tokunaga, "Synthesis of Higher Alcohols by Fischer-Tropsch Synthesis over Alkali Metal-Modified Cobalt Catalysts", *Applied Catalysis A: General*, Vol. 458, pp. 145-154, 2013.

76. Almeida, L. C., O. Sanz, J. D'Olhaberriague, S. Yunes, and M. Montes, "Microchannel Reactor for Fischer-Tropsch Synthesis: Adaptation of a Commercial Unit for Testing Microchannel Blocks", *Fuel*, Vol. 110, pp. 171-177, 2013.

77. Almeida, L. C., F. J. Echave, O. Sanz, M. A. Centeno, G. Arzamendi, L. M. Gandía, E. F. Sousa-Aguiar, J. A. Odriozola, and M. Montes, "Fischer-Tropsch Synthesis in Microchannels", *Chemical Engineering Journal*, Vol. 167, No. 2-3, pp. 536-544, 2011.

78. Jalama, K., N. J. Coville, H. Xiong, D. Hildebrandt, D. Glasser, S. Taylor, A. Carley, J. a. Anderson, and G. J. Hutchings, "A Comparison of Au/Co/Al_2O_3 and Au/Co/SiO_2 Catalysts in the Fischer-Tropsch Reaction", *Applied Catalysis A: General*, Vol. 395, No. 1-2, pp. 1-9, 2011.

79. Ma, W., G. Jacobs, P. Gao, T. Jermwongratanachai, W. D. Shafer, V. R. R. Pendyala, C. H. Yen, J. L. S. Klettlinger, and B. H. Davis, "Fischer-Tropsch Synthesis: Pore Size

and Zr Promotional Effects on the Activity and Selectivity of 25%Co/Al_2O_3 Catalysts", *Applied Catalysis A: General*, Vol. 475, pp. 314-324, 2014.

80. Li, Y., X. Qin, T. Wang, L. Ma, L. Chen, and N. Tsubaki, "Fischer Tropsch Synthesis from H_2-Deficient Biosyngas over Mn Added Co/SiO_2 Catalysts", *Fuel*, Vol. 136, pp. 130-135, 2014.

81. Shimura, K., T. Miyazawa, T. Hanaoka, and S. Hirata, "Fischer-Tropsch Synthesis over TiO_2 Supported Cobalt Catalyst: Effect of TiO_2 Crystal Phase and Metal Ion Loading", *Applied Catalysis A: General*, Vol. 460-461, pp. 8-14, 2013.

82. Jermwongratanachai, T., G. Jacobs, W. Ma, W. D. Shafer, M. K. Gnanamani, P. Gao, B. Kitiyanan, B. H. Davis, J. L. S. Klettlinger, C. H. Yen, D. C. Cronauer, a. J. Kropf, and C. L. Marshall, "Fischer-Tropsch Synthesis: Comparisons between Pt and Ag Promoted Co/Al_2O_3 Catalysts for Reducibility, Local Atomic Structure, Catalytic Activity, and Oxidation-Reduction (or) Cycles", *Applied Catalysis A: General*, Vol. 464-465, pp. 165-180, 2013.

83. De La Osa, A. R., A. De Lucas, J. L. Valverde, A. Romero, I. Monteagudo, P. Coca, and P. Sánchez, "Influence of Alkali Promoters on Synthetic Diesel Production over Co Catalyst", *Catalysis Today*, Vol. 167, No. 1, pp. 96-106, 2011.

84. De La Osa, a. R., A. De Lucas, A. Romero, J. L. Valverde, and P. Sánchez, "Fischer-Tropsch Diesel Production over Calcium-Promoted Co/Alumina Catalyst: Effect of Reaction Conditions", *Fuel*, Vol. 90, No. 5, pp. 1935-1945, 2011.

85. Feyzi, M., M. Irandoust, and A. A. Mirzaei, "Effects of Promoters and Calcination Conditions on the Catalytic Performance of Iron-Manganese Catalysts for Fischer-Tropsch Synthesis", *Fuel Processing Technology*, Vol. 92, No. 5, pp. 1136-1143, 2011.

86. Chu, W., P. A. Chernavskii, L. Gengembre, G. A. Pankina, P. Fongarland, and A. Y. Khodakov, "Cobalt Species in Promoted Cobalt Alumina-Supported Fischer-Tropsch Catalysts", *Journal of Catalysis*, Vol. 252, No. 2, pp. 215-230, 2007.

87. Fu, T., C. Huang, J. Lv, and Z. Li, "Fuel Production through Fischer-Tropsch Synthesis on Carbon Nanotubes Supported Co Catalyst Prepared by Plasma", *Fuel*, Vol. 121, pp. 225-231, 2014.

88. Kang, J., W. Deng, Q. Zhang, and Y. Wang, "Ru Particle Size Effect in Ru/CNT Catalyzed Fischer-Tropsch Synthesis", *Journal of Energy Chemistry*, Vol. 22, No. 2, pp. 321-328, 2013.

89. Gonzalo-Chacón, L., M. Almohalla, E. Gallegos-Suarez, A. Guerrero-Ruiz, and I. Rodríguez-Ramos, "Effects of the Reduction Temperature over Ex-Chloride Ru Fischer-Tropsch Catalysts Supported on High Surface Area Graphite and Promoted by Potassium", *Applied Catalysis A: General*, Vol. 480, pp. 86-92, 2014.

90. Ding, M., Y. Yang, Y. Li, T. Wang, L. Ma, and C. Wu, "Impact of H_2/CO Ratios on Phase and Performance of Mn-Modified Fe-Based Fischer Tropsch Synthesis Catalyst", *Applied Energy*, Vol. 112, pp. 1241-1246, 2013.

91. Kwack, S. H., J. W. Bae, M. J. Park, S. M. Kim, K. S. Ha, and K. W. Jun, "Reaction Modeling on the Phosphorous-Treated Ru/Co/Zr/SiO_2 Fischer-Tropsch Catalyst with the Estimation of Kinetic Parameters and Hydrocarbon Distribution", *Fuel*, Vol. 90, No. 4, pp. 1383-1394, 2011.

92. Yang, J., W. Ma, D. Chen, A. Holmen, and B. H. Davis, "Fischer-Tropsch Synthesis: A Review of the Effect of Co Conversion on Methane Selectivity", *Applied Catalysis A: General*, Vol. 470, pp. 250-260, 2014.

93. Feyzi, M. and A. Akbar Mirzaei, "Performance and Characterization of Iron-Nickel Catalysts for Light Olefin Production", *Journal of Natural Gas Chemistry*, Vol. 19, No. 4, pp. 422-430, 2010.

94. Mirzaei, A. A., A. B. Babaei, M. Galavy, and A. Youssefi, "A Silica Supported Fe-Co Bimetallic Catalyst Prepared by the Sol/Gel Technique: Operating Conditions, Catalytic Properties and Characterization", *Fuel Processing Technology*, Vol. 91, No. 3, pp. 335-347, 2010.

95. Zhou, X., Q. Chen, Y. Tao, and H. Weng, "Correlation between Hydrocarbon Distribution and Water-Hydrocarbon Ratio in Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 20, No. 5, pp. 525-530, 2011.

96. Xing, C., G. Yang, D. Wang, C. Zeng, Y. Jin, R. Yang, Y. Suehiro, and N. Tsubaki, "Controllable Encapsulation of Cobalt Clusters Inside Carbon Nanotubes as Effective Catalysts for Fischer-Tropsch Synthesis", *Catalysis Today*, Vol. 215, pp. 24-28, 2013.

97. Myrstad, R., S. Eri, P. Pfeifer, E. Rytter, and A. Holmen, "Fischer-Tropsch Synthesis in a Microstructured Reactor", *Catalysis Today*, Vol. 147, pp. 3-6, 2009.

98. Holmen, A., H. J. Venvik, R. Myrstad, J. Zhu, and D. Chen, "Monolithic, Microchannel and Carbon Nanofibers/Carbon Felt Reactors for Syngas Conversion by Fischer-Tropsch Synthesis", *Catalysis Today*, Vol. 216, pp. 150-157, 2013.

99. Kantardzic, M., *Data Mining: Concepts, Models, Methods, and Algorithms* Piscataway, New Jersey, IEEE Press, 2003.

100. Jiawei Han, M. K., Jian Pei, *Data Mining: Concepts and Techniques*, 3rd ed, Waltham, Massachusetts: Morgan Kaufmann Publishers, 2012.

101. Tufféry, S., *Data Mining and Statistics for Decision Making*, Wiltshire, United Kingdom: John Wiley & Sons, Ltd, 2011.

102. Kurgan, K., *Data Mining: A Knowledge Discovery Approach*: Springer, 2007.

103. Detroja, K. P., R. D. Gudi, and S. C. Patwardhan, "Plant-Wide Detection and Diagnosis Using Correspondence Analysis", *Control Engineering Practice*, Vol. 15, No. 12, pp. 1468-1483, 2007.

104. Stelle, D., M. C. Barioni, and L. P. Scott, "Using Data Mining to Identify Structural Rules in Proteins", *Applied Mathematics and Computation*, Vol. 218, No. 5, pp. 1997-2004, 2011.

105. Page, R. and R. C. Stevens, "Crystallization Data Mining in Structural Genomics: Using Positive and Negative Results to Optimize Protein Crystallization Screens", *Methods*, Vol. 34, No. 3, pp. 373-389, 2004.

106. Yamasaki, C., K. O. Koyanagi, Y. Fujii, T. Itoh, R. Barrero, T. Tamura, Y. Yamaguchi-Kabata, M. Tanino, J. I. Takeda, S. Fukuchi, S. Miyazaki, N. Nomura, S. Sugano, T. Imanishi, and T. Gojobori, "Investigation of Protein Functions through Data-Mining on Integrated Human Transcriptome Database, H-Invitational Database (H-Invdb)", *Gene*, Vol. 364, No. 1-2, pp. 99-107, 2005.

107. Günay, M. E., I. E. Nikerel, E. Toksoy Oner, B. Kirdar, and R. Yildirim, "Simultaneous Modeling of Enzyme Production and Biomass Growth in Recombinant Escherichia Coli Using Artificial Neural Networks", *Biochemical Engineering Journal*, Vol. 42, No. 3, pp. 329-335, 2008.

108. Simon, L. L. and K. Hungerbuhler, "Industrial Batch Dryer Data Mining Using Intelligent Pattern Classifiers: Neural Network, Neuro-Fuzzy and Takagi-Sugeno Fuzzy Models", *Chemical Engineering Journal*, Vol. 157, No. 2-3, pp. 568-578, 2010.

109. Bansal, S., S. Roy, and F. Larachi, "Support Vector Regression Models for Trickle Bed Reactors", *Chemical Engineering Journal*, Vol. 207-208, pp. 822-831, 2012.

110. Hattori, T. and S. Kito, "Neural Network as a Tool for Catalyst Development", *Catalysis Today*, Vol. 23, No. 4, pp. 347-355, 1995.

111. Sharma, B. K., M. P. Sharma, S. Kumar, S. Kumar, S. B. Tendulkar, and S. S. Tambe, "Fischer-Tropsch Synthesis with $Co/SiO_2–Al_2O_3$ Catalyst and Steady-State Modeling Using Artificial Neural Networks", *Fuel*, Vol. 77, No. 15, pp. 1763-1768, 1998.

112. Adib, H., R. Haghbakhsh, M. Saidi, M. A. Takassi, F. Sharifi, M. Koolivand, M. R. Rahimpour, and S. Keshtkari, "Modeling and Optimization of Fischer-Tropsch Synthesis in the Presence of Co (III)/Al_2O_3 Catalyst Using Artificial Neural Networks and Genetic Algorithm", *Journal of Natural Gas Science and Engineering*, Vol. 10, pp. 14-24, 2013.

113. Zavyalova, U., M. Holena, R. Schlögl, and M. Baerns, "Statistical Analysis of Past Catalytic Data on Oxidative Methane Coupling for New Insights into the Composition of High-Performance Catalysts", *ChemCatChem*, Vol. 3, No. 12, pp. 1935-1947, 2011.

114. Günay, M. E. and R. Yildirim, "Neural Network Analysis of Selective CO Oxidation over Copper-Based Catalysts for Knowledge Extraction from Published Data in the Literature", *Industrial & Engineering Chemistry Research*, Vol. 50, No. 22, pp. 12488-12500, 2011.

115. Odabaşı, Ç., M. E. Günay, and R. Yildirim, "Knowledge Extraction for Water Gas Shift Reaction over Noble Metal Catalysts from Publications in the Literature between 2002 and 2012", *International Journal of Hydrogen Energy*, Vol. 39, No. 11, pp. 5733-5746, 2014.

116. Iida, H., K. Sakamoto, M. Takeuchi, and A. Igarashi, "Fischer Tropsch Synthesis over Co/SiO_2 and Co-M (M: Ru, Re)/SiO_2 Catalysts Prepared by a High-Temperature Super-Critical Drying Method", *Applied Catalysis A: General*, Vol. 466, pp. 256-263, 2013.

117. Prieto, G., P. Concepción, R. Murciano, and A. Martínez, "The Impact of Pre-Reduction Thermal History on the Metal Surface Topology and Site-Catalytic Activity of Co/SiO_2 Fischer-Tropsch Catalysts", *Journal of Catalysis*, Vol. 302, pp. 37-48, 2013.

118. Fu, T., Y. Jiang, J. Lv, and Z. Li, "Effect of Carbon Support on Fischer-Tropsch Synthesis Activity and Product Distribution over Co-Based Catalysts", *Fuel Processing Technology*, Vol. 110, pp. 141-149, 2013.

119. Gnanamani, M. K., G. Jacobs, W. D. Shafer, and B. H. Davis, "Fischer-Tropsch Synthesis: Activity of Metallic Phases of Cobalt Supported on Silica", *Catalysis Today*, Vol. 215, pp. 13-17, 2013.

120. Peña, D., A. Griboval-Constant, V. Lecocq, F. Diehl, and a. Y. Khodakov, "Influence of Operating Conditions in a Continuously Stirred Tank Reactor on the Formation of

Carbon Species on Alumina Supported Cobalt Fischer-Tropsch Catalysts", *Catalysis Today*, Vol. 215, pp. 43-51, 2013.

121. Zeng, B., B. Hou, L. Jia, D. Li, and Y. Sun, "Fischer-Tropsch Synthesis over Different Structured Catalysts: The Effect of Silica Coating onto Nanoparticles", *Journal of Molecular Catalysis A: Chemical*, Vol. 379, pp. 263-268, 2013.

122. De Tymowski, B., Y. Liu, C. Meny, C. Lefèvre, D. Begin, P. Nguyen, C. Pham, D. Edouard, F. Luck, and C. Pham-Huu, "Co-Ru/SiC Impregnated with Ethanol as an Effective Catalyst for the Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 419-420, pp. 31-40, 2012.

123. Park, S. J., S. M. Kim, M. H. Woo, J. W. Bae, K. W. Jun, and K. S. Ha, "Effects of Titanium Impurity on Alumina Surface for the Activity of Co/Ti-Al_2O_3 Fischer-Tropsch Catalyst", *Applied Catalysis A: General*, Vol. 419-420, pp. 148-155, 2012.

124. Bragana, L. F. F. P. G., M. Ojeda, J. L. G. Fierro, and M. I. P. Da Silva, "Bimetallic Co-Fe Nanocrystals Deposited on SBA-15 and HMS Mesoporous Silicas as Catalysts for Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 423-424, pp. 146-153, 2012.

125. Lualdi, M., G. Di Carlo, S. Lögdberg, S. Järås, M. Boutonnet, V. La Parola, L. F. Liotta, G. M. Ingo, and A. M. Venezia, "Effect of Ti and Al Addition via Direct Synthesis to SBA-15 as Support for Cobalt Based Fischer-Tropsch Catalysts", *Applied Catalysis A: General*, Vol. 443-444, pp. 76-86, 2012.

126. Bai, S., C. Huang, J. Lv, and Z. Li, "Comparison of Induction Behavior of Co/CNT and Co/SiO_2 Catalysts for the Fischer-Tropsch Synthesis", *Catalysis Communications*, Vol. 22, pp. 24-27, 2012.

127. Tavasoli, A., S. Karimi, S. Taghavi, Z. Zolfaghari, and H. Amirfirouzkouhi, "Comparing the Deactivation Behaviour of Co/CNT and Co/γ-Al_2O_3 Nano Catalysts in Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 21, No. 5, pp. 605-613, 2012.

128. Jung, J. S., S. W. Kim, and D. J. Moon, "Fischer-Tropsch Synthesis over Cobalt Based Catalyst Supported on Different Mesoporous Silica", *Catalysis Today*, Vol. 185, No. 1, pp. 168-174, 2012.

129. Venezia, A. M., V. La Parola, L. F. Liotta, G. Pantaleo, M. Lualdi, M. Boutonnet, and S. Järås, "Co/SiO_2 Catalysts for Fischer–Tropsch Synthesis; Effect of Co Loading and Support Modification by TiO_2", *Catalysis Today*, Vol. 197, No. 1, pp. 18-23, 2012.

130. Liu, C., J. Li, Y. Zhang, S. Chen, J. Zhu, and K. Liew, "Fischer-Tropsch Synthesis over Cobalt Catalysts Supported on Nanostructured Alumina with Various Morphologies", *Journal of Molecular Catalysis A: Chemical*, Vol. 363-364, pp. 335-342, 2012.

131. Özkara-Aydinoğlu, Ş., Ö. Ataç, Ö. F. Gül, Ş. Kinayyiğit, S. Şal, M. Baranak, and I. Boz, "α-Olefin Selectivity of Fe-Cu-K Catalysts in Fischer-Tropsch Synthesis: Effects of Catalyst Composition and Process Conditions", *Chemical Engineering Journal*, Vol. 181-182, pp. 581-589, 2012.

132. Ma, C., N. Yao, Q. Han, and X. Li, "Synthesis and Application of γ-Al_2O_3 Supported CoRu-Based Fischer–Tropsch Catalyst", *Chemical Engineering Journal*, Vol. 191, pp. 534-540, 2012.

133. Zamani, Y., M. Bakavoli, M. Rahimizadeh, A. Mohajeri, and S. M. Seyedi, "Synergetic Effect of La and Ba Promoters on Nanostructured Iron Catalyst in Fischer-Tropsch Synthesis", *Chinese Journal of Catalysis*, Vol. 33, No. 7, pp. 1119-1124, 2012.

134. Ali, S., N. A. Mohd Zabidi, and D. Subbarao, "Effect of Niobium Promoters on Iron-Based Catalysts for Fischer-Tropsch Reaction", *Journal of Fuel Chemistry and Technology*, Vol. 40, No. 1, pp. 48-53, 2012.

135. Bukur, D. B., Z. Pan, W. Ma, G. Jacobs, and B. H. Davis, "Effect of Co Conversion on the Product Distribution of a Co/Al_2O_3 Fischer–Tropsch Synthesis Catalyst Using a Fixed Bed Reactor", *Catalysis Letters*, Vol. 142, No. 11, pp. 1382-1387, 2012.

136. Cano, L. a., M. V. Cagnoli, J. F. Bengoa, A. M. Alvarez, and S. G. Marchetti, "Effect of the Activation Atmosphere on the Activity of Fe Catalysts Supported on SBA-15 in the Fischer-Tropsch Synthesis", *Journal of Catalysis*, Vol. 278, No. 2, pp. 310-320, 2011.

137. Jia, L., L. Jia, D. Li, B. Hou, J. Wang, and Y. Sun, "Silylated Co/SBA-15 Catalysts for Fischer-Tropsch Synthesis", *Journal of Solid State Chemistry*, Vol. 184, No. 3, pp. 488-493, 2011.

138. Jean-Marie, .A, A. Griboval-Constant, A. Y. Khodakov, and F. Diehl, "Influence of Sub-Stoichiometric Sorbitol Addition Modes on the Structure and Catalytic Performance of Alumina-Supported Cobalt Fischer-Tropsch Catalysts", *Catalysis Today*, Vol. 171, No. 1, pp. 180-185, 2011.

139. Ding, M., Y. Yang, B. Wu, T. Wang, H. Xiang, and Y. Li, "Effect of Reducing Agents on Microstructure and Catalytic Performance of Precipitated Iron-Manganese Catalyst for Fischer-Tropsch Synthesis", *Fuel Processing Technology*, Vol. 92, No. 12, pp. 2353-2359, 2011.

140. Yan, Z., D. B. Bukur, and D. W. Goodman, "Silica-Supported Rhodium-Cobalt Catalysts for Fischer-Tropsch Synthesis", *Catalysis Today*, Vol. 160, No. 1, pp. 39-43, 2011.

141. De La Osa, A. R., A. De Lucas, A. Romero, J. L. Valverde, and P. Sánchez, "Influence of the Catalytic Support on the Industrial Fischer-Tropsch Synthetic Diesel Production", *Catalysis Today*, Vol. 176, No. 1, pp. 298-302, 2011.

142. Zhou, X., Q. Chen, Y. Tao, and H. Weng, "Effect of Vacuum Impregnation on the Performance of Co/SiO_2 Fischer-Tropsch Catalyst", *Journal of Natural Gas Chemistry*, Vol. 20, No. 4, pp. 350-355, 2011.

143. Ali, S., N. A. M. Zabidi, and D. Subbarao, "Development of Niobium-Promoted Cobalt Catalysts on Carbon Nanotubes for Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 20, No. 6, pp. 659-663, 2011.

144. Motchelaho, M. a. M., H. Xiong, M. Moyo, L. L. Jewell, and N. J. Coville, "Effect of Acid Treatment on the Surface of Multiwalled Carbon Nanotubes Prepared from Fe-Co Supported on $CaCO_3$: Correlation with Fischer-Tropsch Catalyst Activity", *Journal of Molecular Catalysis A: Chemical*, Vol. 335, No. 1-2, pp. 189-198, 2011.

145. Park, J. Y., Y. J. Lee, P. R. Karandikar, K. W. Jun, J. W. Bae, and K. S. Ha, "Ru Promoted Cobalt Catalyst on γ-Al2o3 Support: Influence of Pre-Synthesized Nanoparticles on Fischer-Tropsch Reaction", *Journal of Molecular Catalysis A: Chemical*, Vol. 344, No. 1-2, pp. 153-160, 2011.

146. Park, S. J., J. W. Bae, Y. J. Lee, K. S. Ha, K. W. Jun, and P. Karandikar, "Deactivation Behaviors of Pt or Ru Promoted Co/P-Al_2O_3 Catalysts During Slurry-Phase Fischer-Tropsch Synthesis", *Catalysis Communications*, Vol. 12, No. 6, pp. 539-543, 2011.

147. Yao, Y., X. Liu, D. Hildebrandt, and D. Glasser, "Fischer-Tropsch Synthesis Using $H_2/CO/CO_2$ Syngas Mixtures over an Iron Catalyst", *Industrial & Engineering Chemistry Research*, Vol. 50, pp. 11002-11012, 2011.

148. Hong, J., W. Chu, P. a. Chernavskii, and A. Y. Khodakov, "Effects of Zirconia Promotion on the Structure and Performance of Smaller and Larger Pore Silica-Supported Cobalt Catalysts for Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 382, No. 1, pp. 28-35, 2010.

149. Xiong, H., M. Moyo, M. a. M. Motchelaho, L. L. Jewell, and N. J. Coville, "Fischer-Tropsch Synthesis over Model Iron Catalysts Supported on Carbon Spheres: The Effect of Iron Precursor, Support Pretreatment, Catalyst Preparation Method and Promoters", *Applied Catalysis A: General*, Vol. 388, No. 1-2, pp. 168-178, 2010.

150. Nakhaei Pour, A., M. R. Housaindokht, J. Zarkesh, and S. F. Tayyari, "Studies of Carbonaceous Species in Alkali Promoted Iron Catalysts During Fischer-Tropsch Synthesis", *Journal of Industrial and Engineering Chemistry*, Vol. 16, No. 6, pp. 1025-1032, 2010.

151. Kumabe, K., T. Sato, K. Matsumoto, Y. Ishida, and T. Hasegawa, "Production of Hydrocarbons in Fischer-Tropsch Synthesis with Fe-Based Catalyst: Investigations of Primary Kerosene Yield and Carbon Mass Balance", *Fuel*, Vol. 89, No. 8, pp. 2088-2095, 2010.

152. Feltes, T. E., L. Espinosa-Alonso, E. D. Smit, L. D'Souza, R. J. Meyer, B. M. Weckhuysen, and J. R. Regalbuto, "Selective Adsorption of Manganese onto Cobalt for Optimized Mn/Co/TiO_2 Fischer-Tropsch Catalysts", *Journal of Catalysis*, Vol. 270, No. 1, pp. 95-102, 2010.

153. Kang, S. H., J. W. Bae, K. J. Woo, P. S. Sai Prasad, and K. W. Jun, "ZSM-5 Supported Iron Catalysts for Fischer-Tropsch Production of Light Olefin", *Fuel Processing Technology*, Vol. 91, No. 4, pp. 399-403, 2010.

154. Jung, H., J. I. Yang, J. H. Yang, H. T. Lee, D. H. Chun, and H. J. Kim, "Investigation of Fischer-Tropsch Synthesis Performance and Its Intrinsic Reaction Behavior in a Bench Scale Slurry Bubble Column Reactor", *Fuel Processing Technology*, Vol. 91, No. 12, pp. 1839-1844, 2010.

155. Karimi, A., A. N. Pour, F. Torabi, B. Hatami, A. Tavasoli, M. R. Alaei, and M. Irani, "Fischer-Tropsch Synthesis over Ruthenium-Promoted Co/Al_2O_3 Catalyst with Different Reduction Procedures", *Journal of Natural Gas Chemistry*, Vol. 19, No. 5, pp. 503-508, 2010.

156. Bao, A., J. Li, and Y. Zhang, "Effect of Barium on Reducibility and Activity for Cobalt-Based Fischer-Tropsch Synthesis Catalysts", *Journal of Natural Gas Chemistry*, Vol. 19, No. 6, pp. 622-627, 2010.

157. Zhang, C., G. Zhao, K. Liu, Y. Yang, H. Xiang, and Y. Li, "Adsorption and Reaction of CO and Hydrogen on Iron-Based Fischer-Tropsch Synthesis Catalysts", *Journal of Molecular Catalysis A: Chemical*, Vol. 328, No. 1-2, pp. 35-43, 2010.

158. Zhang, Y., J. Bao, S. Nagamori, and N. Tsubaki, "A New and Direct Preparation Method of Iron-Based Bimodal Catalyst and Its Application in Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 352, No. 1-2, pp. 277-281, 2009.

159. Malek Abbaslou, R. M., A. Tavasoli, and A. K. Dalai, "Effect of Pre-Treatment on Physico-Chemical Properties and Stability of Carbon Nanotubes Supported Iron Fischer-Tropsch Catalysts", *Applied Catalysis A: General*, Vol. 355, No. 1-2, pp. 33-41, 2009.

160. Jacobs, G., M. C. Ribeiro, W. Ma, Y. Ji, S. Khalid, P. T. A. Sumodjo, and B. H. Davis, "Group 11 (Cu, Ag, Au) Promotion of 15%Co/Al_2O_3 Fischer-Tropsch Synthesis Catalysts", *Applied Catalysis A: General*, Vol. 361, No. 1-2, pp. 137-151, 2009.

161. Prieto, G., A. Martínez, R. Murciano, and M. a. Arribas, "Cobalt Supported on Morphologically Tailored SBA-15 Mesostructures: The Impact of Pore Length on Metal Dispersion and Catalytic Activity in the Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 367, No. 1-2, pp. 146-156, 2009.

162. Bae, J. W., S. J. Park, S. H. Kang, Y. J. Lee, K. W. Jun, and Y. W. Rhee, "Effect of Cu Content on the Bifunctional Fischer-Tropsch Fe-Cu-K/ZSM-5 Catalyst", *Journal of Industrial and Engineering Chemistry*, Vol. 15, No. 6, pp. 798-802, 2009.

163. Escalona, N., C. Medina, R. García, and P. Reyes, "Fischer Tropsch Reaction from a Mixture Similar to Biosyngas. Influence of Promoters on Surface and Catalytic Properties of Co/SiO_2 Catalysts", *Catalysis Today*, Vol. 143, No. 1-2, pp. 76-79, 2009.

164. Bao, A., K. Liew, and J. Li, "Fischer-Tropsch Synthesis on CaO-Promoted Co/Al_2O_3 Catalysts", *Journal of Molecular Catalysis A: Chemical*, Vol. 304, No. 1-2, pp. 47-51, 2009.

165. Xiong, H., Y. Zhang, K. Liew, and J. Li, "Ruthenium Promotion of Co/SBA-15 Catalysts with High Cobalt Loading for Fischer-Tropsch Synthesis", *Fuel Processing Technology*, Vol. 90, No. 2, pp. 237-246, 2009.

166. Trépanier, M., A. Tavasoli, A. K. Dalai, and N. Abatzoglou, "Fischer-Tropsch Synthesis over Carbon Nanotubes Supported Cobalt Catalysts in a Fixed Bed Reactor: Influence of Acid Treatment", *Fuel Processing Technology*, Vol. 90, No. 3, pp. 367-374, 2009.

167. Liu, Y., T. Hanaoka, T. Miyazawa, K. Murata, K. Okabe, and K. Sakanishi, "Fischer-Tropsch Synthesis in Slurry-Phase Reactors over Mn- and Zr-Modified Co/SiO_2 Catalysts", *Fuel Processing Technology*, Vol. 90, No. 7-8, pp. 901-908, 2009.

168. Sari, A., Y. Zamani, and S. A. Taheri, "Intrinsic Kinetics of Fischer-Tropsch Reactions over an Industrial Co-Ru/γ-Al_2O_3 Catalyst in Slurry Phase Reactor", *Fuel Processing Technology*, Vol. 90, No. 10, pp. 1305-1313, 2009.

169. Tavasoli, A., M. Trépanier, R. M. Malek Abbaslou, A. K. Dalai, and N. Abatzoglou, "Fischer-Tropsch Synthesis on Mono- and Bimetallic Co and Fe Catalysts Supported on Carbon Nanotubes", *Fuel Processing Technology*, Vol. 90, No. 12, pp. 1486-1494, 2009.

170. Hong, J., P. a. Chernavskii, A. Y. Khodakov, and W. Chu, "Effect of Promotion with Ruthenium on the Structure and Catalytic Performance of Mesoporous Silica (Smaller and Larger Pore) Supported Cobalt Fischer-Tropsch Catalysts", *Catalysis Today*, Vol. 140, No. 3-4, pp. 135-141, 2009.

171. Cao, C., J. Hu, S. Li, W. Wilcox, and Y. Wang, "Intensified Fischer-Tropsch Synthesis Process with Microchannel Catalytic Reactors", *Catalysis Today*, Vol. 140, No. 3-4, pp. 149-156, 2009.

172. Ma, X., Q. Sun, W. Ying, and D. Fang, "Effects of the Ratio of Fe to Co over Fe-Co/SiO_2 Bimetallic Catalysts on Their Catalytic Performance for Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 18, No. 2, pp. 232-236, 2009.

173. Ma, X., Q. Sun, W. Ying, and D. Fang, "Effects of Promoters on Catalytic Performance of Fe-Co/SiO_2 Catalyst for Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 18, No. 3, pp. 354-358, 2009.

174. Hao, Q., L. Bai, H. Xiang, and Y. Li, "Activation Pressure Studies with an Iron-Based Catalyst for Slurry Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 18, No. 4, pp. 429-435, 2009.

175. Ding, M., Y. Yanga, B. Wu, J. Xua, C. Zhang, H. Xiang, and Y. Li, "Study of Phase Transformation and Catalytic Performance on Precipitated Iron-Based Catalyst for Fischer-Tropsch Synthesis", *Journal of Molecular Catalysis A: Chemical*, Vol. 303, No. 1-2, pp. 65-71, 2009.

176. Bae, J. W., S. M. Kim, Y. J. Lee, M. J. Lee, and K. W. Jun, "Enhanced Fischer-Tropsch Activity on Co/P-Al_2O_3 Catalyst: Effect of Phosphorous Content", *Catalysis Communications*, Vol. 10, No. 9, pp. 1358-1362, 2009.

177. Li, Y. P., T. J. Wang, C. Z. Wu, X. X. Qin, and N. Tsubaki, "Effect of Ru Addition to Co/SiO_2/HZSM-5 Catalysts on Fischer-Tropsch Synthesis of Gasoline-Range Hydrocarbons", *Catalysis Communications*, Vol. 10, No. 14, pp. 1868-1874, 2009.

178. Jean-Marie, A., A. Griboval-Constant, A. Y. Khodakov, and F. Diehl, "Cobalt Supported on Alumina and Silica-Doped Alumina: Catalyst Structure and Catalytic Performance in Fischer-Tropsch Synthesis", *Comptes Rendus Chimie*, Vol. 12, No. 6-7, pp. 660-667, 2009.

179. Zhao, L., G. Liu, and J. Li, "Effect of La_2O_3 on a Precipitated Iron Catalyst for Fischer-Tropsch Synthesis", *Chinese Journal of Catalysis*, Vol. 30, No. 7, pp. 637-642, 2009.

180. Shan, Y., K. Liew, and J. Li, "Effect of Silylation of SBA-15 on Its Supported Cobalt Catalysts for Fischer-Tropsch Synthesis", *Chinese Journal of Catalysis*, Vol. 30, No. 11, pp. 1091-1095, 2009.

181. Ding, M., Y. Yang, J. Xu, Z. Tao, H. Wang, H. Wang, H. Xiang, and Y. Li, "Effect of Reduction Pressure on Precipitated Potassium Promoted Iron-Manganese Catalyst for Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 345, No. 2, pp. 176-184, 2008.

182. Tavasoli, A., R. M. Malek Abbaslou, and A. K. Dalai, "Deactivation Behavior of Ruthenium Promoted Co/γ-Al_2O_3 Catalysts in Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 346, No. 1-2, pp. 58-64, 2008.

183. Nakhaei Pour, A., S. M. K. Shahri, H. R. Bozorgzadeh, Y. Zamani, A. Tavasoli, and M. A. Marvast, "Effect of Mg, La and Ca Promoters on the Structure and Catalytic Behavior of Iron-Based Catalysts in Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 348, No. 2, pp. 201-208, 2008.

184. Lesaint, C., W. R. Glomm, Ø. Borg, S. Eri, E. Rytter, and G. Øye, "Synthesis and Characterization of Mesoporous Alumina with Large Pore Size and Their Performance in Fischer-Tropsch Synthesis", *Applied Catalysis A: General*, Vol. 351, No. 1, pp. 131-135, 2008.

185. Shi, L., J. Chen, K. Fang, and Y. Sun, "CH3-Modified Co/Ru/SiO_2 Catalysts and the Performances for Fischer-Tropsch Synthesis", *Fuel*, Vol. 87, No. 4-5, pp. 521-526, 2008.

186. Lohitharn, N. and J. G. Goodwin, "Effect of K Promotion of Fe and FeMn Fischer-Tropsch Synthesis Catalysts: Analysis at the Site Level Using SSITKA", *Journal of Catalysis*, Vol. 260, No. 1, pp. 7-16, 2008.

187. Hou, W., B. Wu, Y. Yang, Q. Hao, L. Tian, H. Xiang, and Y. Li, "Effect of SiO_2 Content on Iron-Based Catalysts for Slurry Fischer-Tropsch Synthesis", *Fuel Processing Technology*, Vol. 89, No. 3, pp. 284-291, 2008.

188. Hinchiranan, S., Y. Zhang, S. Nagamori, T. Vitidsant, and N. Tsubaki, "TiO_2 Promoted Co/SiO_2 Catalysts for Fischer-Tropsch Synthesis", *Fuel Processing Technology*, Vol. 89, No. 4, pp. 455-459, 2008.

189. Wang, T., Y. Ding, L. Yuan, H. Zhu, and L. Lin, "Influence of Lanthanum on the Performance of Zr-Co/Activated Carbon Catalysts in Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 17, No. 2, pp. 153-158, 2008.

190. Lira, E., C. M. López, F. Oropeza, M. Bartolini, J. Alvarez, M. Goldwasser, F. L. Linares, J. F. Lamonier, and M. J. Pérez Zurita, "HMS Mesoporous Silica as Cobalt Support for the Fischer-Tropsch Synthesis: Pretreatment, Cobalt Loading and Particle Size Effects", *Journal of Molecular Catalysis A: Chemical*, Vol. 281, No. 1-2, pp. 146-153, 2008.

191. Wan, H., B. Wu, C. Zhang, H. Xiang, and Y. Li, "Promotional Effects of Cu and K on Precipitated Iron-Based Catalysts for Fischer-Tropsch Synthesis", *Journal of Molecular Catalysis A: Chemical*, Vol. 283, No. 1-2, pp. 33-42, 2008.

192. Xiong, H., Y. Zhang, K. Liew, and J. Li, "Fischer-Tropsch Synthesis: The Role of Pore Size for Co/SBA-15 Catalysts", *Journal of Molecular Catalysis A: Chemical*, Vol. 295, No. 1-2, pp. 68-76, 2008.

193. Cai, Q. and J. Li, "Catalytic Properties of the Ru Promoted Co/SBA-15 Catalysts for Fischer-Tropsch Synthesis", *Catalysis Communications*, Vol. 9, No. 10, pp. 2003-2006, 2008.

194. Kim, S. M., J. W. Bae, Y. J. Lee, and K. W. Jun, "Effect of CO_2 in the Feed Stream on the Deactivation of Co/γ-Al_2O_3 Fischer-Tropsch Catalyst", *Catalysis Communications*, Vol. 9, No. 13, pp. 2269-2273, 2008.

195. Mochizuki, T., T. Hara, N. Koizumi, and M. Yamada, "Surface Structure and Fischer-Tropsch Synthesis Activity of Highly Active Co/SiO_2 Catalysts Prepared from the Impregnating Solution Modified with Some Chelating Agents", *Applied Catalysis A: General*, Vol. 317, No. 1, pp. 97-104, 2007.

196. Jalama, K., N. J. Coville, D. Hildebrandt, D. Glasser, and L. L. Jewell, "Effect of Cobalt Carboxylate Precursor Chain Length on Fischer-Tropsch Cobalt/Alumina Catalysts", *Applied Catalysis A: General*, Vol. 326, No. 2, pp. 164-172, 2007.

197. Li, T., Y. Yang, C. Zhang, X. An, H. Wan, Z. Tao, H. Xiang, Y. Li, F. Yi, and B. Xu, "Effect of Manganese on an Iron-Based Fischer-Tropsch Synthesis Catalyst Prepared from Ferrous Sulfate", *Fuel*, Vol. 86, No. 7-8, pp. 921-928, 2007.

198. Borg, Ø., S. Eri, E. a. Blekkan, S. Storsæter, H. Wigum, E. Rytter, and A. Holmen, "Fischer-Tropsch Synthesis over γ-Alumina-Supported Cobalt Catalysts: Effect of Support Variables", *Journal of Catalysis*, Vol. 248, No. 1, pp. 89-100, 2007.

199. Wan, H. J., B. S. Wu, C. H. Zhang, H. W. Xiang, Y. W. Li, B. F. Xu, and F. Yi, "Study on Fe-Al_2O_3 Interaction over Precipitated Iron Catalyst for Fischer-Tropsch Synthesis", *Catalysis Communications*, Vol. 8, No. 10, pp. 1538-1545, 2007.

200. Tristantini, D., S. Lögdberg, B. Gevert, Ø. Borg, and A. Holmen, "The Effect of Synthesis Gas Composition on the Fischer-Tropsch Synthesis over Co/γ-Al_2O_3 and Co-Re/γ-Al_2O_3 Catalysts", *Fuel Processing Technology*, Vol. 88, No. 7, pp. 643-649, 2007.

201. Liu, Y., J. Chen, K. Fang, Y. Wang, and Y. Sun, "A Large Pore-Size Mesoporous Zirconia Supported Cobalt Catalyst with Good Performance in Fischer-Tropsch Synthesis", *Catalysis Communications*, Vol. 8, No. 6, pp. 945-949, 2007.

202. Liu, Y., Y. Zhang, and N. Tsubaki, "The Effect of Acetic Acid Pretreatment for Cobalt Catalysts Prepared from Cobalt Nitrate", *Catalysis Communications*, Vol. 8, No. 5, pp. 773-776, 2007.

203. Nurunnabi, M., K. Murata, K. Okabe, M. Inaba, and I. Takahara, "Effect of Mn Addition on Activity and Resistance to Catalyst Deactivation for Fischer-Tropsch Synthesis over Ru/Al_2O_3 and Ru/SiO_2 Catalysts", *Catalysis Communications*, Vol. 8, No. 10, pp. 1531-1537, 2007.

204. An, X., B. S. Wu, H. J. Wan, T. Z. Li, Z. C. Tao, H. W. Xiang, and Y. W. Li, "Comparative Study of Iron-Based Fischer-Tropsch Synthesis Catalyst Promoted with Potassium or Sodium", *Catalysis Communications*, Vol. 8, No. 12, pp. 1957-1962, 2007.

205. Wan, H. J., B. S. Wu, C. H. Zhang, B. T. Teng, and Z. C. Tao, "Effect of Al_2O_3/SiO_2 Ratio on Iron-Based Catalysts for Fischer-Tropsch Synthesis", *Fuel*, Vol. 85, No. 5, pp. 1371-1377, 2006.

206. Zhang, C. H., Y. Yang, B. T. Teng, T. Z. Li, H. Y. Zheng, H. W. Xiang, and Y. W. Li, "Study of an Iron-Manganese Fischer-Tropsch Synthesis Catalyst Promoted with Copper", *Journal of Catalysis*, Vol. 237, No. 2, pp. 405-415, 2006.

207. Li, H., S. Wang, F. Ling, and J. Li, "Studies on MCM-48 Supported Cobalt Catalyst for Fischer-Tropsch Synthesis", *Journal of Molecular Catalysis A: Chemical*, Vol. 244, No. 1-2, pp. 33-40, 2006.

208. Zhang, Y., K. Hanayama, and N. Tsubaki, "The Surface Modification Effects of Silica Support by Organic Solvents for Fischer-Tropsch Synthesis Catalysts", *Catalysis Communications*, Vol. 7, No. 5, pp. 251-254, 2006.

209. Zhang, C. H., H. J. Wan, Y. Yang, H. W. Xiang, and Y. W. Li, "Study on the Iron-Silica Interaction of a Co-Precipitated Fe/SiO_2 Fischer-Tropsch Synthesis Catalyst", *Catalysis Communications*, Vol. 7, No. 9, pp. 733-738, 2006.

210. Wan, H. J., B. S. Wu, Z. C. Tao, T. Z. Li, X. An, H. W. Xiang, and Y. W. Li, "Study of an Iron-Based Fischer-Tropsch Synthesis Catalyst Incorporated with SiO_2", *Journal of Molecular Catalysis A: Chemical*, Vol. 260, No. 1-2, pp. 255-263, 2006.

211. Tao, Z., Y. Yang, C. Zhang, T. Li, J. Wang, H. Wan, H. Xiang, and Y. Li, "Effect of Calcium Promoter on a Precipitated Iron-Manganese Catalyst for Fischer-Tropsch Synthesis", *Catalysis Communications*, Vol. 7, No. 12, pp. 1061-1066, 2006.

212. Li, H., J. Li, H. Ni, and D. Song, "Studies on Cobalt Catalyst Supported on Silica with Different Pore Size for Fischer-Tropsch Synthesis", *Catalysis Letters*, Vol. 110, No. 1-2, pp. 71-76, 2006.

213. Keyvanloo, K., W. C. Hecker, B. F. Woodfield, and C. H. Bartholomew, "Highly Active and Stable Supported Iron Fischer–Tropsch Catalysts: Effects of Support Properties and SiO_2 Stabilizer on Catalyst Performance", *Journal of Catalysis*, Vol. 319, No. pp. 220-231, 2014.

214. Haghtalab, A. and A. Mosayebi, "Co@Ru Nanoparticle with Core–Shell Structure Supported over γ-Al_2O_3 for Fischer–Tropsch Synthesis", *International Journal of Hydrogen Energy*, Vol. 39, No. 33, pp. 18882-18893, 2014.

215. Das, S. K., S. Majhi, P. Mohanty, and K. K. Pant, "CO-Hydrogenation of Syngas to Fuel Using Silica Supported Fe-Cu-K Catalysts: Effects of Active Components", *Fuel Processing Technology*, Vol. 118, pp. 82-89, 2014.

216. Xing, C., J. Sun, G. Yang, W. Shen, L. Tan, P. Zhu, Q. Wei, J. Li, M. Kyodo, R. Yang, Y. Yoneyama, and N. Tsubaki, "Tunable Isoparaffin and Olefin Synthesis in Fischer–Tropsch Synthesis Achieved by Composite Catalyst", *Fuel Processing Technology*, Vol. 136, pp. 68-72, 2015.

217. Zhao, Y.-H., Y.-J. Wang, Q.-Q. Hao, Z.-T. Liu, and Z.-W. Liu, "Effective Activation of Montmorillonite and Its Application for Fischer-Tropsch Synthesis over Ruthenium Promoted Cobalt", *Fuel Processing Technology*, pp. 2014.

218. Díaz, J. A., A. R. De La Osa, P. Sánchez, A. Romero, and J. L. Valverde, "Influence of CO_2 Co-Feeding on Fischer-Tropsch Fuels Production over Carbon Nanofibers Supported Cobalt Catalyst", *Catalysis Communications*, Vol. 44, pp. 57-61, 2014.

219. Azizi, H. R., A. A. Mirzaei, M. Kaykhaii, and M. Mansouri, "Fischer–Tropsch Synthesis: Studies Effect of Reduction Variables on the Performance of Fe–Ni–Co Catalyst", *Journal of Natural Gas Science and Engineering*, Vol. 18, No. pp. 484-491, 2014.

220. Yao, Y., X. Liu, D. Hildebrandt, and D. Glasser, "Fischer-Tropsch Synthesis Using $H_2/CO/CO_2$ Syngas Mixtures: A Comparison of Paraffin to Olefin Ratios for Iron and Cobalt Based Catalysts", *Applied Catalysis A: General*, Vol. 433-434, No. 1, pp. 58-68, 2012.

221. Wan, H. J., B. S. Wu, T. Z. Li, Z. C. Tao, X. an, H. W. Xiang, and Y. W. Li, "Effects of SiO_2 and Al_2O_3 on Performances of Iron-Based catalysts for Slurry Fischer-Tropsch

Synthesis", *Ranliao Huaxue Xuebao/Journal of Fuel Chemistry and Technology*, Vol. 35, No. 5, pp. 589-594, 2007.

222. Ma, X., Q. Sun, F. Cao, W. Ying, and D. Fang, "Effects of the Different Supports on the Activity and Selectivity of Iron-Cobalt Bimetallic Catalyst for Fischer-Tropsch Synthesis", *Journal of Natural Gas Chemistry*, Vol. 15, No. 1, pp. 335-339, 2006.

223. R Core Team, "R: A Language and Environment for Statistical Computing", R Foundation for Statistical Computing, Vienna, Austria, 2014.

224. RStudio Team, "RStudio: Integrated Development for R", RStudio Inc., Boston, Massachussetts, 2012.

225. Bergmeir, C., J. M. Benitez, "Neural Networks in R Using the Stuttgrat Neural Network Simulator: RSNNS", *Journal of Statistical Software*, Vol. 46, No. 7, pp. 1-26, 2012.

226. Beck, M., "Visualization and Analysis Tools for Neural Networks", 2015, http://CRAN.R-project.org/package=NeuralNetTools, [Accessed June 2015].

227. Garson, G. D., "Interpreting Neural Network Connection Weights", *Artificial Intelligence Expert*, Vol. 6, No. 4, pp. 46-51, 1991.

228. Lek, S., M. Delacoste, P. Baran, I. Dimopoulos, J. Lauga, S. Aulagnier, "Application of Neural Networks to Modelling Nonlinear Relationships in Ecology", *Ecological Modelling*, Vol. 90, pp. 39-52, 1996.

229. Therneau, T., B. Atkinson, B. Ripley, "rpart: Recursive Partitioning and Regression Trees", 2015, http://CRAN.R-project.org/package=rpart, [Accessed June 2015].

230. Liaw, A., M. Wiener, "Classification and Regression by randomForest", *R News*, Vol. 2, No. 3, pp. 18-22, 2002.