University of California, Merced

Modeling the Effect of Polymer and Base Oil Molecule Structure and Chemistry on the Bulk Properties of Lubricants

by

Pawan Panwar

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mechanical Engineering

Committee in charge:

Professor Abel Chuang, Chair
Professor Ashlie Martini, Advisor
Professor Venkattraman Ayyaswamy
Professor James Palko

2022

This dissertation of Pawan Panwar is approved, and it is acceptable in quality and form for publication

Abel Chuang, Chair	Date
Ashlie Martini, Advisor	Date
Venkattraman Ayyaswamy	Date
James Palko	Date

University of California, Merced

I dedicate this dissertation to my loving parents, family members, and friends. It stands as a testament to their belief, unconditional love, and encouragement, which kept my spirits and motivation high during this research pursuit.

Curriculum Vitae

Education

- M.S. Engineering, Milwaukee School of Engineering, 2018
- B.S. Mechanical Engineering, Milwaukee School of Engineering, 2017

Publications

- **Panwar, P.**, Yang, Q., and Martini, A., 2022. PyL3dMD: Python LAMMPS 3D Molecular Descriptors Package. J. Cheminformatics. (Submitted)
- **Panwar, P.**, Yang, Q., and Martini, A., 2022. Temperature-Dependent Density and Viscosity Prediction for Hydrocarbons: Machine Learning and Molecular Dynamics Simulations. J. Chem. Inf. Model. (Submitted)
- **Panwar, P.**, Schweissinger, E., Maier, S., Hilf, S., Sirak, S., and Martini, A., 2022. Effect of Polymer Structure and Chemistry on Viscosity Index, Thickening Efficiency, and Traction Coefficient of Lubricants. J. Mol. Liq.
- Liu. T., **Panwar, P.**, Khajeh, A., Rahman, M. H., Menezes, P. L., and Martini, A., 2022. Review of Molecular Dynamics Simulations of Phosphonium Ionic Liquid Lubricants. Tribol. Lett.
- Khajeh, A., Rahman, M. H., Liu, T., **Panwar, P.**, Menezes, P. L., and Martini, A., 2022. Thermal Decomposition of Phosphonium Salicylate and Phosphonium Benzoate Ionic Liquids. J. Mol. Liq.
- Rahman, M. H., Khajeh, A., **Panwar, P.**, Patel, M., Martini, A., and Menezes, P. L., 2022. Recent Progress on Phosphonium-based room temperature ionic liquids: Synthesis, properties, tribological performances, and applications. Tribol. Int., 167, 107331.
- Johnson, J. L., Garcia-Bravo, J., **Panwar, P.**, and Michael, P., 2021. Strategies to Minimize Data Sample Size for Regression-Based Pump/Motor Models. In Scandinavian International Conference on Fluid Power, pp.134-149.
- **Panwar, P.**, Michael, P., Devlin, M., and Martini, A., 2020. Critical Shear Rate of Polymer-Enhanced Hydraulic Fluids. Lubricants, 8(12), 102.
- **Panwar, P.**, Len, M., Gajghate, N., Michael, P., and Martini, A., 2019. Fluid Effects on Mechanical Efficiency of Hydraulic Pumps: Dynamometer Measurements and Molecular Simulations. In ASME/BATH 2019 Symposium on Fluid Power and Motion Control. American Society of Mechanical Engineers Digital Collection.
- **Panwar, P.** and Michael, P., 2018. Empirical Modeling of Hydraulic Pumps and Motors based upon the Latin Hypercube Sampling Method. Int. J. Hydromechatron, 1(3), pp.272-292.

- Michael, P., Cheekolu, M., **Panwar, P.**, Devlin, M., Davidson, R., Johnson, D., and Martini, A., 2018. Temporary and Permanent Viscosity Loss Correlated to Hydraulic System Performance. Tribol. Trans., 61(5), pp.901-910.
- Johnson, J., Garcia, J., Michael, P., and **Panwar, P.**, 2018. A Sequel - Progressively Sequenced Regression Helps to Establish Minimum Sample Size at Test Time. ISO/TC131/ SC8/WG13/(Math Modeling ad hoc Project), San Antonio, Texas.

Presentations

- Effect of Polymer Structure and Chemistry on Viscosity Index, Thickening Efficiency, and Traction Coefficient of Lubricants, STLE 76th Annual Meeting and Exhibition, Orlando, Florida | 2022.
- Polymer Additives in Lubricants, UC Merced GradSlam, virtually | 2021.
- Critical Shear Rate of Polymer-Enhanced Hydraulic Fluids, STLE 75th Annual Meeting and Exhibition, virtually | 2021.
- Fluid Effects on Mechanical Efficiency of Hydraulic Pumps, STLE Northern California Section, virtually | 2020.
- Polymer-Enhanced Fluid Effects on Mechanical Efficiency of Hydraulic Pumps, STLE 74th Annual Meeting and Exhibition, Minneapolis, Minnesota | 2019.
- Modeling of Axial Piston Pump Input Torque and Output Flow Rate, STLE Chicago Section, Chicago | 2018.
- An Investigation of the Effects of Fluid Composition on Aeration, Efficiency, and Sound Generation in an Axial Piston Pump, CCEFP Summit, Minneapolis, Minnesota | 2018.

Awards

- Graduate Student Researcher grant from the University of California Merced | 2019 - 2022
- Paper was featured on the cover of the recent issue of Lubricants MDPI | 2020
- Received a research grant from STLE Northern California Section | 2020
- Edmond E. Bisson best paper award by STLE | 2020
- Best paper award in ASME/BATH 2019 Symposium on Fluid Power and Motion Control proceeding | 2019
- Graduate tuition grant from MSOE while pursuing M.S. | 2017- 2018
- Our research at Fluid Power Institute MSOE was highlighted with a picture in the Milwaukee Journal Sentinel | 2017
- Received a research grant from STLE Chicago Section | 2017

- Nominated by ME faculty at MSOE for the Outstanding Student Award | 2017
- Academic scholarship received from MSOE while pursuing B.S. | 2015 – 2017
- The senior design project was highlighted in the Dimension magazine of MSOE | 2016

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to express my deep love and affection to my parents. I am deeply indebted to them for their support, motivation, and love throughout my journey. My father, Mr. Laxminarayan Panwar, taught me discipline, manners, respect, and valuable lessons that have helped me succeed in life. My mother, Mrs. Kiran Panwar, most beautiful and kind of all, taught me never to give up. I am grateful for her unconditional love and care. I want to extend my sincere thanks to my sister Sarita, my brother Mr. Radheshyam Panwar and my sister-in-law Mrs. Sarita Panwar for giving me their unequivocal support and encouragement during my tough times. They have always been the epitome of caring family members for my entire life.

I would like to express my deepest gratitude to my advisor and mentor, Prof. Ashlie Martini, for her guidance and encouragement during my Ph.D. work. I have been inspired by her meticulousness, deep research thoughts, scientific knowledge, and approach to shaping students as a confident and independent researcher. She gave me the freedom to do things in my own order and offered me immense opportunities beyond the domain of my Ph.D. work, such as leading the experimental team, being session chair in various conferences, reviewing scientific research articles, etc. This enriched my leadership quality, critical thinking, communication skill, and time management ability and helped me to grow as an independent scientist. Besides my advisor, I thank the rest of my committee members for their insightful comments and questions which incented me to widen my research from various perspectives.

It was a great privilege to work with EVONIK team. Several of my ideas have taken a preliminary shape due to several discussions with this team. It also motivated me to expand my knowledge by learning new tools such as machine learning, quantitative-structure-property-relationship modeling, etc. I thank my fellow Martini lab mates, for their cooperation and helpful suggestions. Working with them and having fun during our meetings, conferences, and parties was a great pleasure. Special thank goes to Dr. Michelle Len for guiding me.

I am grateful to Mr. Paul Michael, who encouraged me to pursue Ph.D. His good advice, support and friendship have been valuable on both an academic and a personal level.

Finally, I would like to express my profound appreciation to my friends Meenakshi, Mubeena, Azhar, and Savannah for their affection, encouragement and for always cheering me up. I enjoyed playing board games, taking adventure trips, watching movies, and having a second round of dinner with them. I also thank adorable little Maisha, who innocently gave me the name 'Pancle'. I cherish all the loving memories shared with my friends!

Abstract

Modeling the Effect of Polymer and Base Oil Molecule Structure and Chemistry on the Bulk Properties of Lubricants

Formulating and designing energy efficient polymer additives and base oil molecules is one way to reduce energy losses in mechanical systems. However, the first step to designing novel polymers and base oils with improved performance under a wider range of operating conditions is to decipher the effect of their structure and chemistry on the bulk properties of lubricants. Therefore, this research aimed to investigate the factors and mechanisms that influence the key bulk properties (density, viscosity, traction coefficient) of liquid lubricants at the molecular level using molecular dynamics (MD) simulations, machine learning (ML), and quantitative-structure-property-relationship (QSPR) modeling.

In the first study, the prospect of improving the mechanical efficiency (ME) of hydraulic systems by formulating fluids with viscosity modifiers (VMs) was tested in a pump dynamometer. Lower viscosity fluids provided better ME but decreasing the viscosity of base oil by adding VM did not have the same effect. Simulations showed that viscosity was directly correlated to elongation of the polymers under shear, which, together with calculations of the key shear rate range in a pump, suggested ways of designing VMs to achieve a specific viscosity profile that maximizes ME. The second study presented a model for predicting the critical shear rate, intending to identify a fluid that shear thins in the key shear rate range in a pump. The model was applied to predict the properties of fluids formulated with VMs and validated by comparison to viscosities obtained from experimental measurements and MD simulations across many decades of shear rates. Results demonstrated that polymer molecular weight plays an important role in determining the critical shear rate, whereas polymer concentration primarily affects the Newtonian viscosity. The simulations showed the molecular origins of shear thinning and critical shear rate.

In the third study, MD simulations were used to identify the QSPR of polymer-enhanced lubricants having commercial grades chemistries. The molecular origins of differences in the viscosity index, thickening efficiency, and traction coefficient between the fluids were investigated by calculating multiple structural properties of the polymers in the simulations. In the fourth study, a python package called PyL3dMD was developed, which enables users to compute nearly 2000 dynamic molecular descriptors by post-processing MD simulation trajectories. This was then used in our fifth study to relate 3D conformations of 305 complex hydrocarbons to their temperature-dependent density and viscosity and, as a result, developed advanced ML-based QSPR models. The models predicted density and dynamic viscosity with the coefficient of determination values of 99.6% and 97.7%, respectively, for all data sets, including a test data set of 45 molecules.

Overall, this dissertation demonstrated the viability of various techniques in understanding molecular interactions and facilitating novel and innovative designs of polymeric additives and base oils for various applications with improved performance.

Chapter 1. Introduction

Tribology is the study and application of the principles of friction, wear, and lubrication of interacting surfaces in relative motion. The word tribology was first derived in 1966 by Peter Jost from the Greek root 'tribo,' which implies 'to rub' or 'to slide,' and the suffix 'logy' which implies 'science of' or 'study of' [1]. Peter Jost introduced tribology as a new field of science to mitigate the great economic impact that wear failures had on the British economy in the mid-20th century. However, despite being a recent field of science, quantitative studies of friction can be traced as far back as 1493, when Leonardo da Vinci first noted the two fundamental 'laws' of friction [2].

In his 1966 report [1], Peter Jost suggested a large-scale implementation of newer and more advanced tribological technologies to save 1.36% of gross annual product (GNP) in the British economy. Similar reports were followed in many major countries such as in Japan (2.6% of the GNP in 1970), in Germany (0.5%, 1976), in the USA (0.79%–0.84%, 1977, 1981) and in China (2%–7%, 1986) [3,4]. Currently, global total annual energy consumption has reached ~440 exajoules (EJ). Out of this, ~23% (119 EJ) originates from tribological contacts. Of that, 20% (103 EJ) is used to overcome friction, and 3% (16 EJ) is used to remanufacture worn parts and spare equipment due to wear and wear-related failures [4,5].

By taking advantage of the new surface, materials, and lubrication technologies for friction reduction and wear protection, energy losses due to friction and wear could potentially be reduced by 40% in the long term (15 years) and by 18% in the short term (8 years) globally [4,5]. Particularly in the USA, 2.1% of the GNP could be saved by introducing advantaged technology realized by targeted research initiatives in tribology [6]. By reducing friction and wear, which are at the core of improving fuel economy and reducing greenhouse gas emissions, we will save monetarily and help limit global warming since 81.5% of the global primary energy is extracted from fossil fuels [4,5]. Lubricants are the primary means of minimizing metal-to-metal contact, therefore reducing friction and wear. So, it is vital to understand how lubricant reduces friction that accounts for approximately a third of the world's primary energy consumption and mitigates breakdown of lubricants and wear losses that cause 70% of equipment failures [7].

1.1. Lubrication

Lubricants are substances that help reduce friction and wear by introducing a physical barrier between interacting surfaces in relative motion. Lubricants can be solids, liquids, and gases, with liquids being the most common and widely used form of lubricants. Solid and gas lubricants are typically used in applications where liquid lubricants have limited functionality. For example, gas lubricants, such as nitrogen and helium, are used where the film thickness is ultra-small and can carry a very low load [8]. On the other hand, solid lubricants, such as graphite and molybdenum disulfide, are used in extreme environmental conditions and carry higher loads. Conditions where solid lubricants are used, include temperatures above 500 °C, vacuum, and outer space where conventional liquid lubricants are less effective [9]. Liquid lubricants typically come in the form of lubricating oils, which are a blend of base oil and additives [8]. For lubricants to function properly in wide range of harsh environments, base fluids need the help of chemicals, called additives.

The most important function of the lubricants is to form a film between surfaces in relative motion to minimize metal-on-metal contact and reduce the force necessary to move one surface against the other, thereby reducing wear and saving energy [8,10,11]. A lubricant also performs many functions that ultimately help protect and prolong the life of the equipment. Some of these functions include cooling, cleaning, protecting, and transferring power [8,10]. In hydraulic systems, lubricants are used to transfer high density power [8,12] in addition to its normal function of lubrication. In some machines, lubricants are also used as a cooling agent, such as in hydraulic machines and combustion engines carry heat from the source of heat generation due to friction to heat sink/radiator. This way, lubricants also help dissipate heat away from critical moving parts, which decreases the possibility of component deformation and wear [8,13]. Lubricants help suspend and remove carbon, sludge, varnish, dirt, and wear debris to facilitate the smooth operation of the equipment [8,13]. Lubricants prevent damage caused by oxidation, corrosion, and wear. By forming a physical film that coats the metal surface, the lubricant prevents the surface from being susceptible to oxygen, water, and acids [8,10,13]. Some lubricants have rust and corrosion inhibitors, extreme-pressure additives, and anti-wear agents that form physical or chemical protective coatings on metal surfaces [8].

1.2. Lubrication Regime

Knowledge of the lubrication regimes in which our machines run is essential to choose the optimum viscosity and type of lubricant with the target to avoid wear and improve energy saving; that is why the Stribeck curve is needed. The thickness

of a lubricating film is used to characterize lubricated interfaces into three different regimes, as illustrated by the Stribeck curve in Figure 1.1. A lubricant's effectiveness in these regimes is quantified by change in the friction coefficient, which is affected by changes in lubricant viscosity, speed, and load. For example, the lubricating film will be too thin at low speed, low viscosity, and high load, leading to high friction. On the other hand, at high speed, high viscosity, and low load, forming a thicker film will reduce friction caused by surface-surface contact. As the viscosity of the lubricant increases, the lubricating film get thicker, as a result, an increase in the coefficient of friction and viscous drag is observed.

Based on the thickness of the lubricating film and the mechanism of film formation [11], the lubrication regimes are divided into boundary lubrication, mixed-film lubrication, and full-film lubrication.

1. The boundary lubrication regime occurs at low speed, low viscosity, or high load, where a fluid film of sufficient thickness cannot form. Additives, such as anti-wear, friction modifiers, and extreme pressure additives, in lubrication provide a molecularly thin protective layer in the boundary regime which form when additives physically or chemically attach onto the contacting surfaces [14].

2. The mixed lubrication regime is considered to exist when the contact between two rough surfaces is partially reduced by a film of pressurized liquid [10]. Systems that fall in this regime have interfaces that experience both boundary and full-film lubrication. Thus, the load-carrying capacity is divided between contacting surfaces and a lubricating film.

3. The full-film lubrication regime occurs at high speed, high viscosity, or low load, where a sufficient thickness fluid film of pressurized liquid completely separates the interacting surfaces [11]. The full-film lubrication regime can further be subcategorized as elastohydrodynamic lubrication (EHL) and hydrodynamic lubrication (HL) regimes. The shapes of the interfacing surfaces often used to classify these subcategories. The contact between highly loaded counter-formal surfaces is classified as the elastohydrodynamic regime [15,16], and contact between conformal surfaces is termed hydrodynamic lubrication [11]. Counter-formal surfaces do not fit perfectly together; therefore, the load is ununiformly distributed along the surface, and as a result, the contact areas elastically deform, and fluid viscosity is high enough to form a load-bearing lubricating film [11,17]. In contrast, in conformal surface no surface-surface contact except for the startup and stop or elastic deformation is observed and the load is distributed evenly along the surface because conformal surfaces geometrically fit well together [11,17].

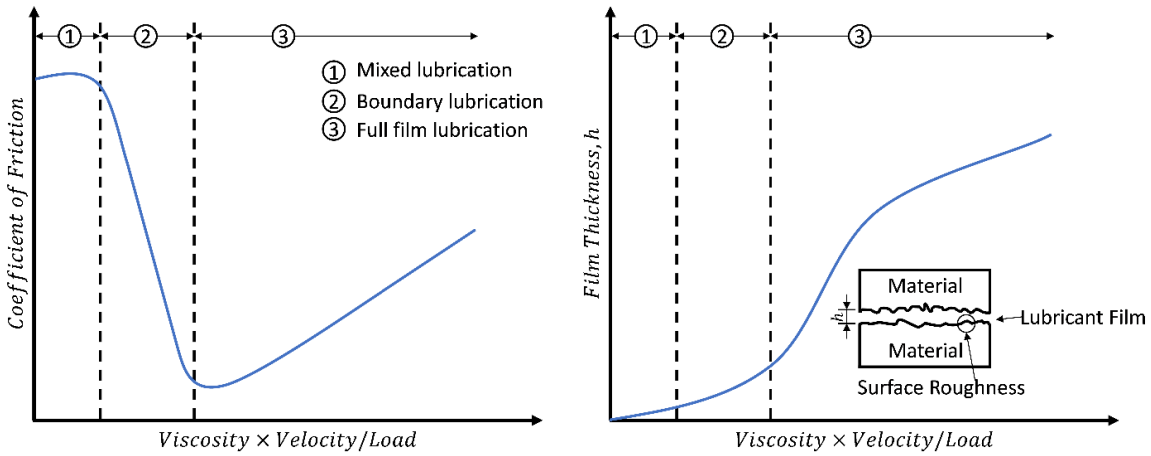


Figure 1.1 The Stribeck curves that shows the various lubrication regimes: (a) describes the relationship between friction and viscosity, load, and speed, and (b) describes the relationship between film thickness and viscosity, load, and speed.

In a lubricated contact, surface roughness can quantitively determine the lubrication regime based on its magnitude relative to the lubricant film thickness. This can be quantified by the so-called lambda ratio, as defined [18,19] by Equation (1.1).

$$\lambda = h_{min} \Big/ \sqrt{R_{a,\,b}^2 + R_{a,J}^2} \quad \begin{cases} \text{boundary lubrication} & \text{if } \lambda < 1 \\ \text{mixed lubrication} & \text{if } 1 < \lambda < 3 \\ \text{full film lubrication} & \text{if } \lambda > 3 \end{cases} \tag{1.1}$$

Here, $R_{a,b}$ and $R_{a,J}$ are the average or RMS roughness of the two surfaces in contact, and h_{min} is the minimum film thickness between them.

1.3. Bulk Properties of Lubricants

The effectiveness of lubricating films is highly dependent on their physical and chemical properties, which are affected by the composition of the lubricant [8,10,11]. For example, in the full film lubrication regime, lubricant viscosity and density play critical roles in application functionality. Similarly, the chemical properties of a lubricant are important in boundary lubrication applications, where chemical interactions between lubricant and surface enable the formation of a lubricating film. Therefore, minimization of friction in the previously mentioned regimes is achieved by improving the properties of the lubricating fluids. Bulk properties vary with machines operating conditions such as pressure (load), temperature, and speed (shear rate).

1.3.1. Density

Density, a measure of mass per unit volume, is a physical property that plays an important role in lubricant function and machine design [10]. Density increases with pressure and decreases with temperature. As fluid density increases, contaminant suspension time also increases. An increase in suspension time can be disastrous to the machine components, depending on the system used. For example, an increased contaminant suspension time could lead to corrosion or cavitation in systems sensitive to contaminants such as hydraulic systems [20]. In high-velocity or turbulent systems, the flow of these fluids can lead to corrosion and wear of surfaces. The knowledge of a lubricant's density, or specific gravity, is important because it indicates the energy required to pump it [8].

1.3.2. Viscosity

Viscosity is a measure of internal resistance to the flow. It is the most important property of a lubricant. There are two types of viscosity – dynamic and kinematic. Dynamic viscosity is the tangential force per unit area required to move one horizontal plane with respect to another plane - at a unit velocity - when maintaining a unit distance apart in the fluid. Kinematic viscosity is the ratio of dynamic viscosity to the density that is a quantity in which no force is involved. Viscosity plays a direct and significant role in forming a lubricating film and ultimately determines a component's success or failure [10]. A higher viscosity fluid leads to a thicker film and higher resistance to the flow (higher viscous friction) or vice versa. The viscosity of lubricants varies with temperature, pressure, and shear rate [21–24]. As pressure or temperature increases, viscosity exponentially increases or decreases, respectively. The viscosity of the fluid decreases by nearly 50% for every 10°C increase in temperature. The viscosity roughly doubles for every 50 MPa increase in pressure [25]. The high-pressure viscosity of the fluid is particularly important in elastohydrodynamic gear lube and rolling contact bearing lubrication [11]. As the shear rate increases, the viscosity of a dilatant fluid increases and a pseudoplastic fluid decreases [10]. Most of the lubricants are pseudoplastic in nature.

1.3.3. Traction Coefficient

Like viscosity, traction (or viscous friction) is a complex but important pressure- and shear-dependent fluid property that quantifies the fluid's resistance to shear. Traction is due entirely to the fluid properties and so is not to be confused with contact friction, which is associated with surface interactions. The traction response is dominated by the shear behavior of the lubricant in the central high contact pressure region of contact, and this

property of a fluid is quantified by a parameter known as traction coefficient (TC). Generally, the traction coefficient increases with increasing slide-to-roll ration [26–28] and contact pressure [26,27], and decreasing entrainment speed [12,27,29,30] and decreasing temperature [26,27].

1.3.4. Other Properties

Other lubricant properties such as specific heat, thermal conductivity, surface tension, bulk modulus, thermal expansion coefficient, oxidation resistance, electrical conductivity, thermal diffusivity, flash point, volatility, demulsibility, and oxidation stability, are also important in the lubrications [8]. Understanding the effects of these features is important to application-specific lubricant selection. However, these properties are not in the scope of this dissertation.

1.4. Lubricant Composition

All modern lubricants are formulated using a base fluid and a performance package, and in the case of the multigrade oils, an additional viscosity modifier. Unformulated lubricants (pure mineral base oils and synthetic base stocks) do not possess the necessary properties to perform effectively in today's highly demanding lubricating environments. Additives improve the lubricating ability of the base oils, either by enhancing the desirable properties already present in the base oils and base stocks or by adding new properties. Hence, additives are an integral part of modern lubricants [8]. Usually, modern lubricants consist of 65-98% base oils, 0-30% viscosity modifier, and 2-20% performance package, depends on the desired performance level and the severity of the end-use requirements [8]. Since the major fraction of a lubricant is a base oil, the bulk properties of base oil have a significant impact on the bulk properties a lubricant. Therefore, when formulating a lubricant, a great deal of importance must be given to the base oil selection.

1.4.1. Base Oils

The bulk properties of base oils are directly related to their source, structure, chemistry, and degree of refinement [1, 5, 8]. Based on the refinement and chemical content, the American Petroleum Institute (API) classifies base oil and stocks for lubricants into five categories listed in Table 1.1 with their distinctive physical properties. The first three are mineral-based base oils, fourth is polyalphaolefins (PAO) synthetic base oils, and fifth is for a base oil that neither falls in the category of mineral and synthetic base oils. Mineral oils are refined crude oil products and contain paraffin, naphthene, and aromatics to varying degrees, which pose lubricant formulation challenges

[34]. Although, as you move from API Group I to API Group III, the degree of refinement increases; as a result, most of the properties (oxidation/thermal stability, volatility, low-temperature performance, traction efficiency, and viscosity) of base oil improve except for the solvency. Synthetic base oils are artificial products manufactured mostly from petroleum-derived low-molecular-weight raw materials through various chemical reactions [1]. These products have well-defined structures and properties. Examples of synthetic base fluids include PAO and polyisobutylene (PIB). The key performance advantage of synthetic oils is their extreme low- and high-temperature performance in certain applications, which cannot be achieved using non-synthetics.

Table 1.1 API classification of base oils and their properties.

Parameters	Group I (GI)	Group II (GII)	Group III (GIII)	Group IV (GIV)	Group V (GV)
Saturates, %	< 90	≥ 90	≥ 90	Polyalphaolefins (PAOs)	Others
Sulfur, ppm	> 0.03	≤ 0.03	≤ 0.03		
Viscosity 100°C, cSt	4–32	4–30	4–8	4.0–7.6	4–70
Viscosity Index (VI)	$\geq 80, < 120$	$\geq 80, < 120$	> 120	> 120	> 100
Pour Point, °C	−15	−15	−15	−22.5 to -15	-45

1.4.2. Additives

Additives are chemicals added in small quantities to the base oil to enhance or introduce new properties to the base fluid. Additives perform three roles in any lubricant: (1) enhance the desirable properties of the base oil, (2) suppress the undesirable ones, and (3) add new properties to the lubricant that improve its overall qualities [31]. These chemicals can improve the friction and wear characteristics of the oil, provide oxidative resistance, control corrosion and contamination, reduce the pour point, inhibit the generation of foam, as well as reduce the decline of viscosity with temperature [8,10]. Lubricant characteristics, such as viscosity, cleanliness of the oil, and degradation of the oil through oxidation, can be controlled by adding additives that specifically improve bulk fluid properties and extend the lifetime of the oil. Additives, such as viscosity modifiers, pour point depressants, detergents, antioxidants, anti-foam agents, emulsifiers, and dispersants, are used to improve viscosity, control chemical breakdown, and control contamination of lubricating oil.

Out of all these additives, viscosity index improvers and viscosity modifiers are the most used additives in the applications where efficiency matters. Viscosity index improvers (VIIs) a specific class of VMs added to the base oils to increase the viscosity index, that is, to increase the broader temperature performance capability of the base oils. They functionally improve the temperature-viscosity

relationship of oil, where viscosity modifiers are thickeners that improve the low and high-temperature functionality of oil. In contrast, viscosity index improvers maintain the low-temperature viscosity but enhance the high-temperature functionality of oil. Film-forming additives are additives that specifically improve lubricity and protect metal surfaces. These additives include antiwear additives, friction modifiers, extreme pressure additives, and corrosion inhibitors. Surface additives function by physically or chemically attaching themselves to surfaces to provide a protective or lubricating layer [13,32,33].

1.5. Dependence of Efficiency on the Bulk Properties

Hydraulic systems function and perform tasks through a pressurized fluid, which is controlled directly or automatically by control valves, distributed through hoses, displaced by pumps, and actuated by cylinders and motors. The efficiency of hydraulic power transmission is primarily affected by internal leakage flow in pumps and motors and by friction and viscous drag in pumps, motors, and cylinders. These losses generate heat and reduce the power available to engage the payload [34,35]. The overall efficiency of hydraulic machines can therefore be separated into two components: volumetric efficiency associated with internal leakage, and mechanical efficiency due to frictional losses. The product of the volumetric and mechanical efficiencies is the overall efficiency of a hydraulic system or component such as a pump or motor [35].

Both volumetric and mechanical efficiency depend on the viscosity of the hydraulic fluid since the fluid functions as both lubricant and power transmission medium. However, fluid viscosity has opposite effects on the two types of efficiency. Specifically, volumetric efficiency increases with increasing viscosity because thicker fluids leak less; in contrast, mechanical efficiency decreases with increasing viscosity because thicker fluids exhibit higher viscous friction and drag. Therefore, the overall efficiency of hydraulic systems varies with fluid viscosity non-monotonically [35–40].

The effect of fluid viscosity on efficiency is further complicated by the fact that viscosity itself depends on operating conditions. Notably, the viscosity of hydraulic fluids decreases rapidly as system temperature increases [41], an effect that is more significant in applications with small oil reservoirs and compact heat exchangers, which lead to higher operating temperatures. Because of this, many hydraulic fluids are formulated with viscosity modifiers (VMs) or viscosity index (VI) improvers to reduce the variation of viscosity during machine operation [42,43]. These additives are long-chain polymers that incorporate repeating molecular blocks of acrylate, olefin, styrene, ethylene, butylene, and alkyleneglycol monomers. Fluids with VMs flow easily at

low temperatures and mitigate heat-related viscosity loss at high temperatures, improving the fluid's performance as a lubricant at both low and high temperatures [44].

It has been proposed that viscosity modifiers (VMs) positively impact efficiency through three mechanisms: (i) reducing leakage flow by thickening, (ii) reducing viscous friction through beneficial shear thinning, and (iii) reducing traction via formulation with lower molecular weight base oils. First, VMs increase viscosity at high temperatures, which decreases leakage. The most frequently cited mechanisms to explain thickening at high temperatures are coil expansion and polymer-polymer association [43]. High-temperature thickening has been found to reduce pressure driven flow losses, thus improving the volumetric efficiency [40,45]. Second, it has been proposed that VMs decrease the viscosity of fluids with increasing shear rate by means of shear thinning. VM polymers may be subject to high shear in areas such as the piston-bore interface in pumps and the roller-cam interface in radial piston motors. Irreversible or permanent viscosity loss results from mechanical scission of polymer molecules [43]. High viscosity index fluids that exhibited low permanent viscosity loss in sonic shear tests were found to provide enhanced pump volumetric efficiency [46]. It has also been shown that temporary viscosity loss has a more significant effect than permanent viscosity loss on leakage flow [3]. However, the third potential benefit of VMs - to reduce traction by enabling the use of lower viscosity base fluid - has not been directly tested.

An axial piston pump and a radial piston motor efficiency tests on five commercial ISO 46 viscosity grade fluids revealed that a low traction fluid decreased the low-speed torque losses as much as by 50% [30]. Similar findings were presented [12] in which a low traction coefficient fluid was found to improve overall low-speed efficiency of a hydraulic motor by 5-16% via transitioning out of the boundary lubrication region at a lower Stribeck number. As a result, low traction is preferable in hydraulic fluids [12,29,30] to decrease shearing forces. In contrast, high traction is preferable in continuously variable transmissions fluids to avoid slippage while minimizing contact friction for energy efficiency [47–49]. Good traction characteristics in a lubricants are essential to achieving a balance between film thickness and resistance to shear to ensure both wear protection and energy efficiency [50].

1.6. Functionality and Mechanisms of Polymeric Additives

Additives are vital for the proper lubrication and prolonged use. Without many of these, the oil would become contaminated, break down, leak out, or not properly protect for wide range of machines operating conditions such as

temperature, pressure, and shear rate. In this research, we only study polymer additives that are used to improve viscosity when added to base oil and their positive and negatives effect on bulk properties of the lubricants. The term polymer additives will be used to refer to VMs.

1.6.1. Viscosity Index and Thickening Efficiency

The viscosity of lubricants, an important fluid property that quantifies the resistance to flow, decreases rapidly as temperature increases [41,51,52]. This is a major concern for lubricants in applications that are operated at a wide range of temperatures; viscosity loss is a particular issue in applications with small oil reservoirs and compact heat exchangers. If the viscosity of a lubricant is too low at a high temperature, a thin lubricating film may no longer have adequate load carrying capacity, leading to asperity contact. The increase in asperity contact increases mechanical friction, which causes machine inefficiency, and surface wear which ultimately leads to component failure [32,53]. However, if a higher viscosity fluid is used to ensure a thick lubricating film of adequate load carrying capacity at high temperatures, that might lead to a poor efficiency at low temperatures due to increase in viscous friction. Therefore, it is desirable for lubricant viscosity to decrease as slowly as possible with increasing temperature. Viscosity index (VI) is the most used metric to quantify the viscosity–temperature relationship of a lubricant, as defined by ASTM D2270 standard [54]. Another important metric is thickening efficiency (TE) that describes the amount of polymer (that is, polymer treat rate) that required in a lubricant formulation to reach either a desired kinematic viscosity or dynamic viscosity at a temperature or 100°C [55]. It is desirable to use as little polymer as possible to achieve good temperature-viscosity behavior, i.e. high TE and VI [56].

Fluids are formulated with polymeric additives to optimize the viscosity-temperature and traction behavior [10,26,61,62,28,42–44,57–60]. Polymers in lubricants are used to serve range of purposes and perform their function through a variety of mechanisms. The molecular weight of polymers ranges from 1-100 kg/mol [10,44,58]. Typical polymeric additives include polyalpha olefin (PAO), olefin copolymer (OCP), polyalkyl methacrylate (PAMA), polyisobutylene (PIB), styrene block copolymers (SBCP), and hydrogenated styrene–diene (HSD) [43,44,57,63]. All these polymers will increase viscosity relative to the viscosity of base oil but might have varying effects on the rate of change of viscosity with temperature for the same backbone length, depending on their composition and architecture. The mechanisms by which polymers perform their function depend on chemistry [64–66] and architecture [55,67,68] both of the polymer and the base oil.

The effect of molecular structure, architecture, and chemistry of polymeric additives on viscosity derived properties, e.g. VI and TE, have been explored using experimental methods [62,69–73] and molecular dynamics (MD) simulation tools [43,59,79,61,62,72,74–78]. MD simulation is a computer simulation tool that is used to study the interactions between chemical species at the atomic scale. The most often reported mechanism by which polymers increase VI of the fluids is coil expansion at higher temperature [80]. In fact, viscosity can be correlated to coil size quantified by the root-mean-square radius of gyration of the polymer [81,82]. While the coil expansion mechanism is widely accepted in the literature [10,44,55,57,58,83,84], it has been observed in some but not in all polymers both experimentally [51,72,85–87] and atomistic simulations [61,72,76,77]. For example, the coil size of some polymers, such as OCP and hydrogenated diene copolymers, remains constant or decreases with increasing temperature, whereas the coil size of polymers derived from esters, such as PAMA, increases with temperature [51,85–87]. Therefore, in general, hydrocarbon-based polymers such as OCP, PIB, PAO, SBCP, and HSD are known to provide better TE, while ester-based polymers such as PAMAs provide better VI [8,55]. It has also been suggested that polymers that expand with temperature might increase VI more than those that do not expand [51]. For example, a block linear styrene–butadiene polymer configuration exhibited the least change in viscosity with temperature compared to alternating and random linear styrene–butadiene polymers, due to ability of the block structure to form smaller coils with more intramolecular interactions at lower temperatures and then expand at higher temperatures [77]. A MD simulation study of PMA and OCP polymers revealed that the presence of electronegative atoms, such as the oxygen in methacrylate, is a key factor in determining if a polymer will expand or contract with temperature [76]. A similar finding was reported in a MD study where PAMA and star PAMA were found to have the highest VI due to polar backbones in contrast to the nonpolar backbones in linear OCP and hyperbranched HBPE polymers [72].

The molecular weight and molecular weight distribution of polymers are easily accessible but important parameters in characterizing polymers. TE was found to increase linearly with the percentage of molecular weight in the backbone of the polymer [57]. It has been also suggested that high molecular weight linear polymers provide better TE [57]. Higher ethylene content OCP polymers have better TE while PAMAs have higher VI [57]. Knowing that the distribution of molecular weight plays a significant role in characterizing rheological properties of fluids, various studies investigated the effect of chain shape (such as linear, comb, and star) [69,72,73,75,88]. An experimental study on polyethylene revealed that an increase of polymer molecular weight by

intermolecular crosslinking significantly improved TE, which ultimately reduced the treat rate[67,68]. On the other hand, another study [69] suggested that viscosity drastically decreased on addition of a hyperbranched polymer instead of a linear polymer. By narrowing the distribution of molecular weight of the polymer (that is, the ratio of mass average molar mass to the number average molar mass) excellent TE can be achieved [57,89].

1.6.2. Shear Thinning

While VMs are needed for operation across a range of temperatures, they also change the rheological properties of fluids, introducing more complex viscosity behavior that is a function of shear stress. At low shear stress conditions, fluids are near their equilibrium state and have viscosity that is independent of shear rate (or stress), i.e., the Newtonian viscosity. However, the high shear stresses in lubricating gaps can cause reversible and nonreversible viscosity loss, also called shear thinning. Reversible, or temporary, viscosity loss occurs when polymers elongate and align with the shear plane but return to their original configuration when shear forces are removed. Irreversible, or permanent, viscosity loss results from mechanical scission of polymer molecules, resulting in lower molecular weight molecules [40,90]. Both types of shear thinning are functions of the concentration, size, and architecture of the polymer molecules [67,72,91–94].

Studies conducted on hydraulic fluids in a dynamometer have suggested that temporary viscosity loss occurs when the fluids are exposed to high shear rates in streamline flows, whereas permanent viscosity loss occurs when fluids are exposed to intense and prolonged shear rates under more extreme conditions [40,95]. Therefore, when formulating hydraulic fluids, VM polymers are chosen such that they do not mechanically degrade and form smaller, less effective polymer fragments when subjected to very high shear rates [44,96,97]. However, fluids will still experience temporary viscosity loss. In fact, even small molecules may exhibit shear thinning in the elastohydrodynamic lubrication regime because of progressive alignment of molecules with shear rate [98,99]. Rheological measurements of low molecular weight squalane in both the Newtonian and non-Newtonian regimes suggested that shear thinning increases with increasing pressure and decreasing temperature [99]. Temporary shear thinning behavior is often described by models with power-law scaling of shear stress with shear rate [11]. Commonly used equations to capture this behavior in lubricants and hydraulic fluids are the Carreau, Cross, and Ree–Eyring models [11]. Such models predict that viscosity will be constant at low shear rates ($\eta = \eta_0$) and then decrease rapidly ($\eta = f(\dot{\gamma})$) above some critical shear rate $\dot{\gamma}_{cr}$. The critical shear rate characterizes the transition from Newtonian to non-Newtonian (shear thinning) behavior. This critical value is

usually taken to be the shear rate at which the tangent to the power law regime intersects with the Newtonian viscosity line [11].

1.6.3. Traction Coefficient

Polymer additives can increase [26,28] or decrease [28] the TC of a lubricant, depending on formulation. TC is typically characterized experimentally using a Mini-Traction-Machine (MTM), which is a ball-on-disk tribometer, at a range of operating conditions. Such studies showed that when the polyethylene polymer concentration was increased in mineral base oil from 0.0 to 2.0%, the TC increased [26]. In that study, a semi-empirical model was also developed to determine the effect of operating conditions such as pressure, temperature, and polymer concentration on TC [26]. The model showed that TC is a function of the natural logarithm of shear rate [26].

Experimental study showed that higher naphthenic carbon content of mineral oil correlates to higher shear resistance [100], some compounds with a gem-dimethyl structure showed higher TC than others [101], and the maximum TC of aliphatic hydrocarbon oils increased with increasing degree of branching [102]. An investigation reported that the flow activation volumes, which was calculated by applying Eyring's hole theory, for viscous flow of 26 lubricating oils was linearly related to the maximum TC [103]. Another study [104,105] investigated 37 molecular structures of traction base fluids of various chemical groups (naphthene, aromatic, paraffin, ether, ketone, alcohol, chloride, and amide) to understand the factors influencing traction and recommended that high molecular stiffness, large size, short alkylene chain length, high melting point, low molecular polarity for the molecule structure to provide a high TC. Polar compounds have a high viscosity due to the molecular interaction force under semi-static conditions but have a low TC due to the repulsion of negative charges under traction conditions (that is, high pressure and high shear rate) [106]. Molecular mechanics modelling [107] qualitatively correlated the TC of ester compounds to the angle of rotation around the C-C bond between ester groups and the potential energy. Other simulations showed that compounds with saturated rings have a high TC and molecular stiffness governs traction properties [108].

A MD study suggested that a polycyclic compounds with ester functional groups are preferable for reducing traction [109] due to slip between rings. Although such studies indicate that simulations can be used to understanding the molecular mechanisms TC, the few studies [109–114] that have been performed either focused on method development or understanding TC only for very small molecules.

1.7. Dissertation Outline

MD simulations have been widely utilized to simulate various properties of mostly fluids of small molecules. However, most of the commercially available lubricants have large polymeric additives. The relationship between molecular properties and the performance metrics and properties of the lubricant is still not fully understood, inhibiting design of new fluids with potentially improved performance. In this work, in addition to accurately modeling bulk properties of commercially available lubricants of large molecules, we seek to fundamentally understand how polymeric additive and base oil molecule structure and chemistry affect the bulk properties of lubricants. These key performance bulk properties of lubricants include density, viscosity, shear viscosity, viscosity index, thickening efficiency, and traction coefficient. Then, extending these findings to developing simple empirical and machine learning based QSPR models to predict these bulk properties via considerable smaller MD simulations.

To summarize, this work aims to study the effect of polymer and base oil structure and chemistry on the bulk properties of the lubricant via quantitative-structure-property-relationship (QSPR) using experimental measurements, molecular dynamics simulations, and machine learning. Chapter 1 introduced the importance of polymeric additives and the various roles that polymeric additives serve to modulate the properties of liquid lubricants. The next five chapters introduce and demonstrate the techniques that were used, and report results obtained in the research. Chapter 2 describes an investigation of the effect of lower molecular weight (1300 g/mol) polymeric additive on mechanical efficiency of hydraulic pumps via dynamometer measurements and molecular simulations. Chapter 3 focuses on laying the groundwork for design of fluids with tunable viscosity behavior and optimized hydraulic efficiency via correlating critical shear rate (onset of shear thinning) and exploring the molecular origins of the shear thinning behavior. Chapter 4 focuses on simulating the Newtonian viscosity and traction coefficient of commercially available polymer-enhanced fluids and then modeling the effect of polymer structure and chemistry on viscosity index, thickening efficiency, and traction coefficient of similar viscosity fluids using the best subset multiple regression analysis. Chapter 5 focuses on developing a one-of-a-kind python package PyL3dMD for calculating thousands of 3D features of the molecules by post-processing MD simulations data. These 3D features encode the chemical and physical information of the molecules which then be used to establish QSPR. Chapter 6 focuses on establishing QSPR relationship between temperature dependent density and viscosity of 305 complex hydrocarbons with their 2D and 3D features, which are calculated

using PyL3dMD and MD simulations, via machine learning. Chapter 6 also focuses on developing a detailed machine learning approach for developing simple and easily interpretable QSPR models. Finally, Chapter 7, summarizes the outcomes of Chapter 2 to Chapter 6. In addition, building on these outcomes, we propose future possible research work to further understand the effects of chemistry and structure of lubricant molecules on their key performance properties. Overall, this dissertation developed various techniques in understanding molecular interactions with bulk properties of lubricants and prepared ground for facilitating design of novel polymers and base oils with improved performance.

Chapter 2. Energy Efficient Hydraulic Fluids - Is Shear Thinning be Beneficial?

2.1. Introduction

In hydraulic fluid power systems, the rotary mechanical power produced by an electric motor or engine is converted into fluid power by a positive displacement pump. The kinetic energy (flow) produced by the pump is transmitted by the hydraulic fluid media. Hydraulic control valves direct the fluid to cylinders and motors that actuate machine functions and convert fluid power back into mechanical power. The interaction between the machine functions and fluid medium stores potential energy within the pressure envelop of the hydraulic system. Internal leakage, migration of fluids from high-pressure zones to low-pressure zones inside the hydraulic system, reduces the amount of power that a machine can deliver. Similarly, both viscous and boundary friction inside the hydraulic displacement machines increases the amount of power that a machine requires to actuate. The desire to maximize efficiency drives many of the design concepts in fluid power technology with the goal of simultaneously reducing internal leakage at high temperatures and viscous friction at low temperatures. Since both leakage and friction are dependent on the viscosity of the hydraulic fluid, one approach to improving both is the use of polymeric viscosity modifiers (VMs) that thicken the fluids. Most hydraulic fluids used in off-highway vehicles are formulated with VMs to mitigate rapid viscosity loss with increasing temperature. It had been assumed that viscous friction can be reduced in machines operate at high shear, such as hydraulic pumps, by enabling shear thinning to our benefit by formulating fluids with VMs, however, this has never been evaluated.

In this study, we isolated the role of VMs in reducing torque losses by studying fluids with similar viscosity profiles but differing VM concentration and base fluid viscosity. Specifically, four fluids were formulated by mixing low traction synthetic polyalphaolefin (PAO) base oils and polyisobutylene (PIB) viscosity modifiers or bis(2-ethylhexyl) adipate synthetic ester (DEHA). The ratio of PAO to PIB polymer in three of the blends was controlled such that the overall viscosity of the solutions was the same at 40°C. The rheology of the fluids was measured, and then pump dynamometer tests were carried out to determine the effects of the fluid composition on pump performance. The dynamometer tests were conducted under low-shear idle and high-shear full power conditions per the ISO 4409 standard method. Pump performance was characterized for the four different fluids in terms of input torque, output flow, volumetric, and mechanical efficiency for a wide range of operating conditions. In addition to this, the dynamometer results were analyzed using pump models. Shear rates

at the critical lubricating gaps in the pump were estimated and pump performance results discussed in the context of viscosity trends at these shear rates. Lastly, molecular dynamics simulations were used to model the four fluids and analyze correlations between the polymer's response to shear and solution viscosity.

2.2. Materials and Methods

2.2.1. Test Fluids

The four hydraulic fluids evaluated in this study are listed in Table 2.1. All the fluids were formulated with an antiwear additive and PAO base oils. The naming convention for PAO is based upon the kinematic viscosity at 100°C. Hence, the test fluids were formulated with PAO of different viscosity, as shown in Table 2.1. HV46-1 and HV46-2 incorporated varying levels of PIB polymer (1,300 g/mol) in 2 cSt and 4 cSt PAO respectively. Both HV46-3 and HV15 do not contain polymer additives. HV46-3 was formulated with PAO 8 whereas HV15 was formulated with PAO 4 and a synthetic adipate ester.

Table 2.1 Chemical composition and properties of the test fluids.

Fluid ID	HV46-1	HV46-2	HV46-3	HV15
PAO 2	61.5%			
PAO 4		81.0%		80.0%
PAO 8			100.0%	
Polyisobutylene	38.5%	19.0%		
Bis(2-ethylhexyl) adipate				20.0%
Kin. Vis. at 40°C, cSt	48.92	46.75	46.74	14.2
Kin. Vis. at 100°C, cSt	8.889	8.075	7.858	3.51
Viscosity Index	164	146	138	128
Specific gravity at 15.6°C	0.8358	0.8345	0.8327	0.8520
Thermal exp. at 20°C, 1/°C	6.34E-4	6.32E-4	6.36E-4	6.77E-4
Flash Point, °C	162	222	258	216
Shear Stability at 40°C	0.84%	0.62%	0.36%	-
Sheared Vis. at 40°C, cSt	48.51	46.46	46.57	-
Sheared Vis. at 100°C, cSt	8.816	8.038	7.873	-

2.2.2. Fluid Properties

Several methods were used to evaluate the properties of the test fluids. The low shear rate viscosity of the fluids was measured via the ASTM D445 Standard Test Method for Kinematic Viscosity of Transparent and Opaque Liquids [115]. The viscosity index, specific gravity, and flash point of test fluids

were determined using ASTM D2270 [54], ASTM D4052 [116], and ASTM D92 [117] test standards, respectively. Permanent viscosity loss was measured via the ASTM D5621 Standard Test Method for Sonic Shear Stability of Hydraulic Fluids [96]. Kinematic viscosities were measured at 40°C and 100°C after sonic shearing. These standard fluid properties are listed in Table 2.1.

The dynamic viscosity of the fluids was measured under low shear rate conditions by conducting multiple viscometry tests in a Cannon StressTech HR Oscillatory Rheometer. The shear rates varied from 20 to 500 1/s, with 10 measurements taken at logarithmically spaced intervals in that range. This procedure was repeated three times for each fluid to report average viscosity values at low shear rates. Thus, Newtonian behavior was evaluated. High shear viscosities were measured using a PCS Ultra Shear viscometer at shear rates from 500,000/s to 2,000,000/s.

2.2.3. Pump Testing

Pump testing was conducted via a modified ISO 4409 test procedure. ISO 4409 "Hydraulic fluid power – Positive displacement pumps, motors, and integral transmissions – Methods of testing and presenting basic steady-state performance" specifies circuit schematics, instrument accuracy ranges, test procedures, and reporting requirements for characterizing positive displacement machine performance [118]. The circuit schematic is shown in Figure 2.1.

An open-circuit axial piston pump with precision electronic displacement control was employed in this study. The power to drive the pump was supplied by a variable frequency drive and a 75HP inverter-type electric motor. The electric motor was instrumented with a torque transducer and rotary encoder. Rotational frequency, inlet oil temperature, swash plate angle, and pump outlet pressure were controlled via automation. System pressure was regulated by a pilot-operated cartridge valve. The pump outlet flow rate was measured using a gear-type positive displacement flow meter positioned upstream from the pressure control valve. The hydraulic fluid was cooled and filtered in the return line. The rotational frequency of the pump ranged from 600 to 2200 rpm, the displacement ranged from 0 to 100%, and the system pressure was varied from 7 to 207 Bar. The nominal fluid temperature at pump inlet was at 50°C. One unique aspect of the test rig was that it incorporated a Coriolis flow meter between the oil reservoir and the pump inlet. This made possible the real-time measurement of fluid density at the inlet pressure.

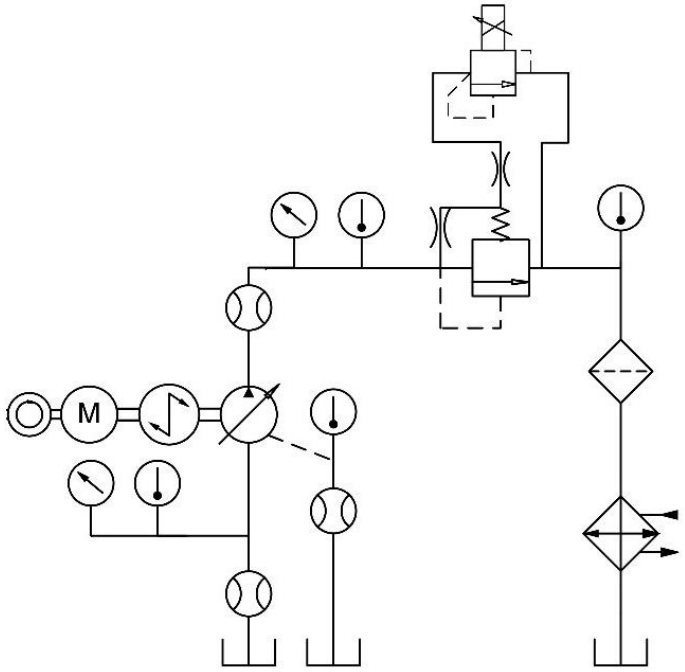


Figure 2.1 Circuit schematic for the hydraulic pump test (test Rig at FPI, MSOE).

2.2.4. Molecular Dynamics Simulation

Molecular dynamics (MD) simulations were used to model the test fluids. Previous gas chromatography-mass spectroscopy measurements reported that PAO 2 consisted of nearly 100% decene dimer $C_{20}H_{42}$, PAO 8 was predominantly decene tetramer $C_{40}H_{82}$, and PAO 4 consisted of equal parts decene trimer and tetramer [119]. The PIB polymer had a nominal molecular weight of 1300 g/mol and a reported bromine number of 12, indicating it was a linear olefin with five unsaturated bonds ($C_{98}H_{188}$) and terminated by exo and endo groups [120]. Models of the PAO base oils, PIB polymer, and DEHA ester were created using Material Studio software, as shown in Figure 2.2.

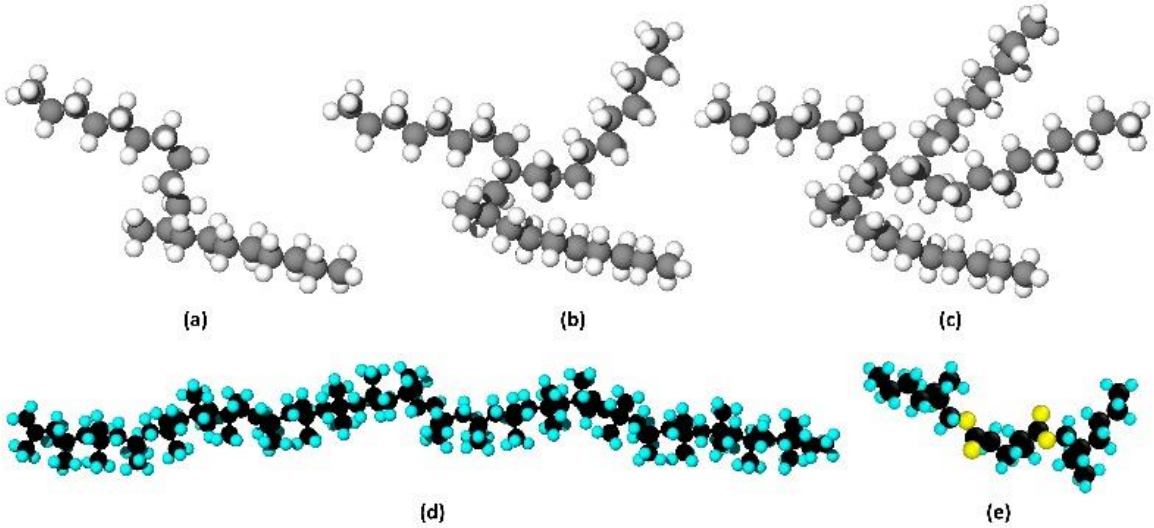


Figure 2.2 Molecular structures of (a) decene dimer, (b) decene trimer, (c) decene tetramer, (d) PIB, and (e) DEHA. The white and blue spheres represent hydrogen atoms, the gray and black spheres represent carbon atoms, and the yellow spheres represent oxygen atoms.

The test fluids were created by placing the appropriate numbers of solute and solvent molecules in a simulation box with periodic boundaries and an approximate aspect ratio of 1:1:2. The number of solvent and solute molecules was determined based on the percent weight concentration. For instance, the HV46-1 model consisted of 4 PIB molecules and 31 decene dimers (PAO 2 cSt) in a 2.3 nm x 2.3 nm x 4.6 nm simulation box with periodic boundary conditions. Similarly, fluids HV46-2, HV46-3, and HV15 were modeled by placing the respective number of PIB polymer, DEHA ester, and PAO base oil molecules in an appropriate box size as listed in Table 2.2. The model systems are shown in Figure 2.3.

Table 2.2 Simulation plan for each fluid.

Molecule	HV46-1	HV46-2	HV46-3	HV15
$C_{20}H_{42}$	31	0	0	0
$C_{60}H_{62}$	0	25	0	28
$C_{40}H_{82}$	0	19	48	21
$C_{98}H_{188}$	4	4	0	0
$C_{22}H_{42}O_4$	0	0	0	16
Box Size, nm^3	2.3 × 2.3 × 4.6	2.8 × 2.8 × 5.6	2.3 × 2.3 × 4.6	2.8 × 2.8 × 5.8

Molecular dynamics simulations were run using Large Atomic/Molecular Massively Parallel Simulation (LAMMPS) software [121] with All Atom Optimized Potentials for Liquid Simulations Force Field for Long Hydrocarbons [122,123]. This force field was modified to better predict density

for hydrocarbons that consist of more than 16 carbon atoms [124]. The temperature and pressure were controlled in the simulation using the Nosé–Hoover thermostat and barostat [125,126], with damping coefficients of 100 and 1000 fs, respectively.

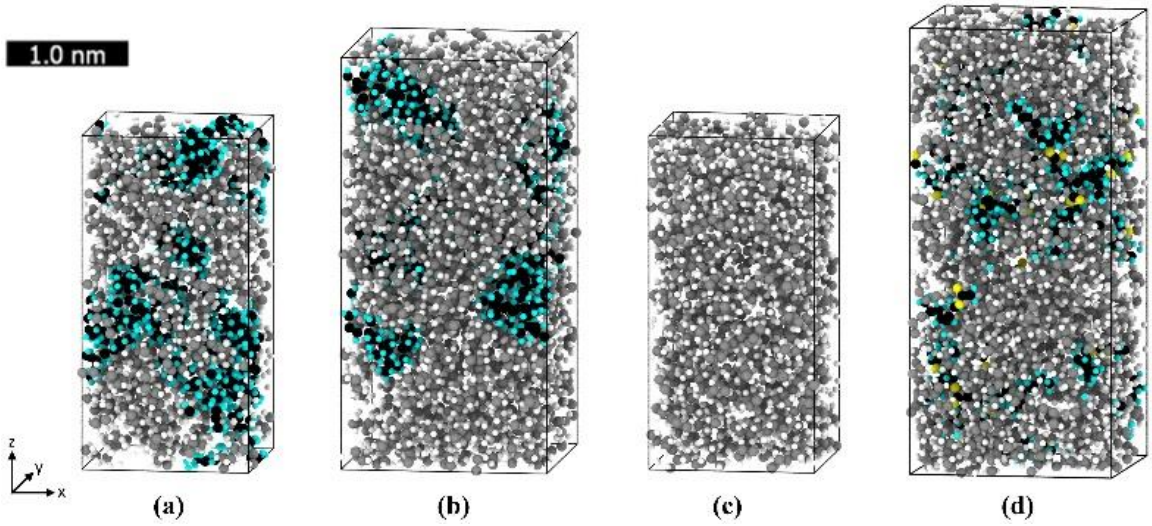


Figure 2.3 Initial configurations of the models of (a) HV46-1, (b) HV46-2, (c) HV46-3, and (d) HV15 that consist of PAO base stock and PIB or DEHA ester as reported in Table 2.2, where the solid black lines of the simulation box indicate the periodic boundaries.

First, the potential energy of each model was minimized, followed by equilibration using a constant number of atoms, volume, and temperature (NVT) at a high temperature for 0.5 ns. Then, the systems were equilibrated using a constant number of atoms, pressure, and temperature (NPT) at 1 atm for 2.0 ns at 50 °C. This was repeated until the density at the last timestep matched the average density of the system. Then, each system was again equilibrated under NVT at 50 °C for approximately 0.5 ns. Viscosity was calculated using non-equilibrium simulations (NEMD) [127–130]. In this method, shear was imposed in the *x*-direction for 100-200ns, during which the shear stresses and velocities of the atoms were calculated. The shear strain rate was calculated from the slope of a linear fit to the average atom velocities, and it was assumed that NEMD simulation reached steady state if the goodness of the linear fit was 99% or greater. Viscosity was calculated from the ratio of the average shear stress to the average shear strain rate. This process was repeated for each test fluid at 50°C for shear rates ranging from 10^9 to 10^{11} 1/s.

2.3. Results and Discussion

2.3.1. Dynamometer Tests

The pump performance tests were conducted for rotational frequencies from 600 to 2200 rpm, displacements from 20% to 100%, system pressures from 41 to 207 bar, and a nominal fluid temperature of 50°C. The test points were selected using the Latin Hypercube sampling method [131]. A statistical description of the torque and flow characteristics of each fluid is reported in Table 2.3. The sample size for each fluid was the same (N=138), enabling a direct comparison of the statistics. These test conditions covered the full working range of the pump, except for idle.

Table 2.3 Pump mean torque and flow for the full operating range.

Fluid	Torque Input [N·m]	Torque Loss [N·m]	Discharge Flow [lpm]	Leakage Flow [lpm]
HV46-1	60.2 ± 3.1	5.1 ± 0.2	35.7 ± 2.0	3.2 ± 0.1
HV46-2	60.1 ± 3.1	5.0 ± 0.2	35.8 ± 2.0	3.1 ± 0.1
HV46-3	60.9 ± 3.1	5.5 ± 0.2	35.8 ± 2.0	3.1 ± 0.1
HV15	58.3 ± 3.0	3.4 ± 0.1	34.7 ± 2.0	4.1 ± 0.1

The standard errors of the mean (SEM) were the same for torque and flow, irrespective of the fluid. The mean torque for the test fluids ranged from 58.3 to 60.9 N·m; a difference of 2.6 N·m. The SEM was approximately 3.0 N·m. Likewise the range of the mean discharge flow rates was less than the SEM of the flow rate. Therefore, the differences between the fluids are statistically insignificant. The 95% confidence intervals for the mean torque and flow rate are shown in Figure 2.4(a). The interval bars are wide, but the differences between the fluids are negligible due to the test conditions, rather than experimental error. Differences in fluid performance and shear response will become evident when low and high-pressure pump performance are evaluated independently.

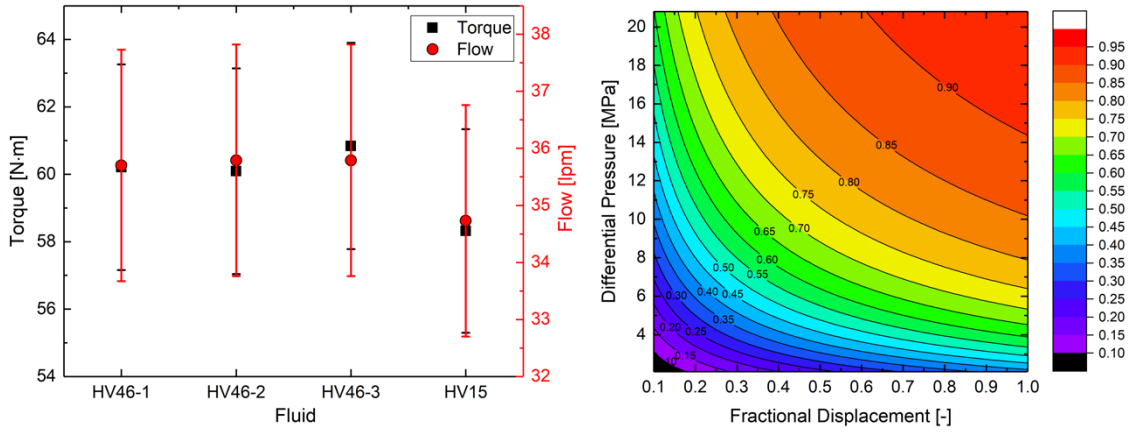


Figure 2.4 Pump mean torque and flow for wide-operating range at 50°C inlet oil temperature. (b) Contour plot of pump mechanical efficiency as a function of differential pressure and fractional displacement (data from FPI, MSOE).

The mechanical efficiency of a pump is defined as the ratio of theoretical torque (T_i) to effective or measured torque (T_e) [132]. Theoretical torque is a function of the pump displacement and differential pressure. The effective torque is the sum of the theoretical torque and the torque required to overcome friction. Pump friction tends to be fairly constant at low speeds and gradually increase with output power. Since friction increases slowly, hydraulic mechanical efficiency increases as the pump output power increases. This effect is illustrated in Figure 2.4(b) where the pump mechanical efficiency (including idle conditions) is plotted as a function of pressure and displacement. It must be noted that at low pressures and displacements, such as experienced during idling conditions, the pump mechanical efficiency is very low.

It is important to consider machine duty cycles and applications when optimizing lubricant performance. Many hydraulic machines in the construction, agriculture, and material handling applications spend long periods of time idling. Under low pressure or idle conditions, hydraulic systems that incorporate variable displacement pumps maintain low-pressure, low-displacement conditions to provide ready power for machine actuation while minimizing fuel consumption. Each fluid was evaluated under low-pressure and low-displacement pump conditions to compare idle losses. As shown in Figure 2.5(a), HV46-1, HV46-2, and HV46-3 produced similar idle torque losses. The idle torque losses for HV15 were 30% lower. This is directly correlated to the lower kinematic viscosity of the HV15 fluid. Because the pump outlet pressures were low under idle conditions, the shear rates in the critical lubricating gaps were in the range of 10^5 1/s. Hence, shear-thinning, and base oil effects were not expected to significantly affect torque.

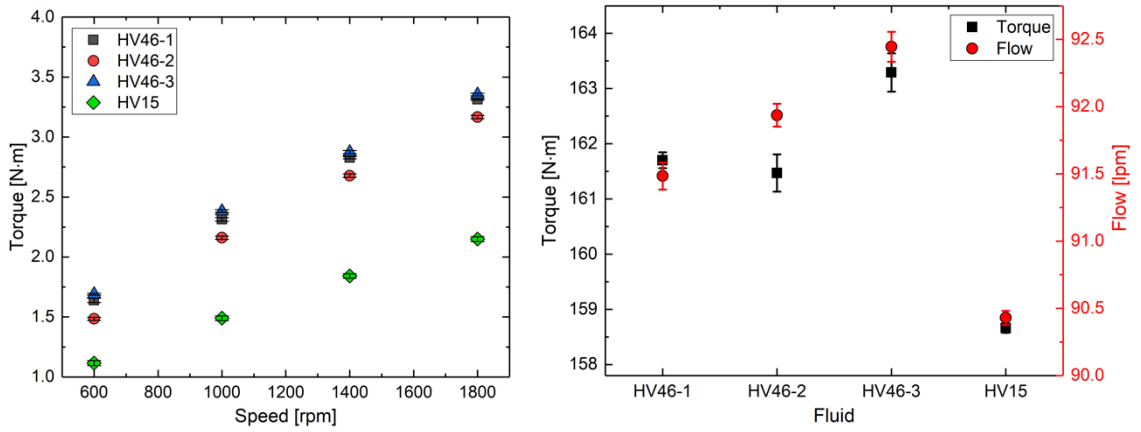


Figure 2.5 (a) Pump torque under idle conditions at 50°C inlet oil temperature and 0.7 MPa (100 psi) outlet pressure with the electronic displacement control commanded to fully Off-stroke. (b) Pump mean torque and flow at 3000 psi, 2200 rpm, 100% displacement, 50°C inlet oil temperature (data from FPI, MSOE).

In high-intensity fluid power applications, the hydraulic system pressures and flow rates are elevated. Under these conditions, energy use and shear rates are at their highest values. Pump performance data were collected for a rotational frequency of 2200 rpm, displacement of 100%, system pressure of 207 bar (3000 psi), and a nominal fluid temperature of 50°C. As shown in Figure 2.5(b), HV46-3 exhibited the highest torque, while the torque values for HV46-1 and HV46-2 were slightly lower. It is hypothesized that this difference (although small) is due to the PIB that was only present in those two fluids. Once again HV15 exhibited the lowest torque.

2.3.2. Pump Modeling

Empirical models for pump flow and torque were created to study the effects of fluid properties on pump performance. The empirical models used physics-based terms incorporating pump operating conditions (rotational frequency, differential pressure, and percent displacement) and fluid properties (density, bulk modulus, and viscosity). The model was described in more detail previously [131]. The density of the test fluids was directly measured at the pump inlet temperature using a Coriolis flow meter. The bulk modulus was calculated as a function of density and pressure based upon the measurements of Kiran [119]. The kinematic viscosity of each fluid at the pump inlet temperature was calculated from the viscosity of the fresh fluid (listed in Table 2.1) using the Walther equation [133]. The kinematic viscosity (υ) values thus derived were multiplied by the density (ρ) of the fluid at the pump inlet temperature to determine the dynamic viscosity (μ).

In the model selection process, three criteria were used for assessing the model quality: the mean standard error (S), the coefficient of determination (R^2), and the Variance Inflation Factor (VIF). The mean standard error is the mean sum of squares of the difference between experimentally determined flow losses, and the flow losses predicted by the model. The mean standard error is useful because it expresses the model error in the units of the dependent variable. The coefficient of determination is the proportion of the variation in the dependent variable (flow loss) that is predictable from the independent variables (displacement, rotational frequency, and differential pressure). The VIF is useful for screening models to prohibit multicollinearity. Multicollinearity occurs when there is a high correlation between model terms. This is undesirable because multicollinearity results in models that over-fit data. Models that over-fit data can generate high residuals when applied to new or independent data sets [134]. To avoid collinear terms, the MATLAB code that was used to select model terms and coefficients eliminated models that yielded high VIF values. The model selection process identified a three-term model for pump input torque that had a high coefficient of determination, a low standard error, and a low VIF. The model included theoretical torque (T_i), viscous torque losses (T_l), and turbulent torque losses (T_t). These terms can respectively be expressed by Equations (2.1), (2.2), (2.3).

$$T_i = \frac{(V_i)p}{2\pi} \tag{2.1}$$

$$T_l = \mu\omega(V_i) \tag{2.2}$$

$$T_t = \frac{\rho(V_i)^{\frac{5}{3}}}{4\pi}\omega^2 \tag{2.3}$$

Thus, input torque was regressed using the above coefficients combined into the expression shown in Equation (2.4).

$$T = C_o + C_1\frac{(V_i)p}{2\pi} + C_2\mu\omega(V_i) + C_3\frac{\rho(V_i)^{\frac{5}{3}}}{4\pi}\omega^2 \tag{2.4}$$

The mean standard error of the torque model was 0.64 N·m with an R^2 value of 99.97% and a VIF less than 3. Since the model has a high R^2 value and a low standard error, a plot of measured versus modeled torque values produces a straight line with minimal scatter. In regression analysis, the difference between the observed value and the modeled value is called the residual. Since the torque model includes terms for pressure, speed, displacement, density, and low-shear kinematic viscosity, systematic errors observed in the residual analysis are due to other factors, such as shear thinning or base oil effects. In order to probe for shear thinning and base oil effects, torque model residuals were plotted as a function of the measured torque, as shown in Figure 2.6(a).

HV15 was not included in this analysis because it had a lower fresh oil viscosity. It was found that the HV46-3 torque model residuals were higher, or more positive, than HV46-1 and HV46-2. The residuals for HV46-1 and HV46-2 were very similar. Since HV46-1 was formulated with PAO 2 and HV46-2 was formulated with PAO 4, residual analysis indicates that differences in diluent oil viscosity did not affect pump torque.

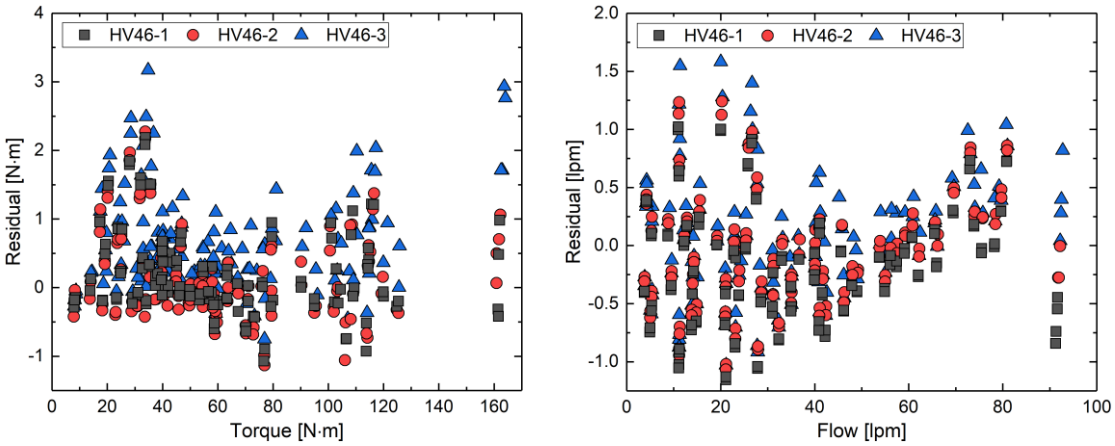


Figure 2.6 (a) Model residuals for the 3-term pump torque model. (b) Residuals for the 3-term flow model.

One potential explanation for the higher torque model residuals produced by HV46-3 is that the use of this fluid increased the pump flow rate. An empirical flow model was developed to compare residuals. The model selection process identified a three-term model for the pump output flow rate that had a low standard error and VIF. The model included terms for theoretical flow rate (Q_i), compressible flow losses (Q_k), and pressure driven flow losses (Q_l). These terms can respectively be expressed by Equations (2.5), (2.6), (2.7).

$$Q_i = V_i\omega \tag{2.5}$$

$$Q_k = \frac{\omega p}{K} \tag{2.6}$$

$$Q_l = \frac{\Delta p}{\mu\omega} \tag{2.7}$$

Thus, output flow was regressed using the above coefficients combined into the expression shown in Equation (2.8).

$$Q = C_o + C_1(V_i\omega) + C_2\left(\frac{\omega p}{K}\right) + C_3\left(\frac{\Delta p}{\mu\omega}\right) \tag{2.8}$$

The standard error of the flow model was 0.46 lpm with an R^2 value of 99.97% and a VIF less than 2. A plot of the flow model residues versus the measured pump flow rate is shown in Figure 2.6(b). Once again, the HV46-3 flow model residuals were higher, or more positive, than HV46-1 and HV46-2. The residuals for HV46-1 and HV46-2 were very similar. Hence, the higher effective torque levels for HV46-3 shown in the Figure 2.6(a) residuals plot appear to be the result of a slightly higher pump flow rate. The key observation from residual analysis was that using a low molecular weight polymeric VM to enable formulation with lower viscosity base oil did not provide beneficial shear thinning. The small reductions in torque that were observed were offset by a similar reduction in the pump output flow.

2.3.3. Towards Tunable Polymer Performance

The results reported in the previous sections indicated that VM polymers are unlikely to improve mechanical efficiency by enabling the use of a lower viscosity base fluid and that, instead, the overall solution viscosity is more important. However, this finding might be due to the fact that the polymers are small and so may not exhibit significant shear thinning at the shear rates relevant to the dynamometer tests. To test this hypothesis, the first step is necessarily to determine the relevant shear rates within the pump. Friction losses occur in axial piston pumps within lubricating gaps, which serve as a sliding bearing between surfaces that move relative to each other and providing a sealing function to control internal leakage. There are three major lubricating gaps in an axial piston pump: piston/cylinder, slipper/swash plate, and cylinder/valve plate, as indicated in purple in Figure 2.7. Gap heights are not a constant. They are affected by the operating pressure, rotational frequency, viscosity, and shear-response of the fluid.

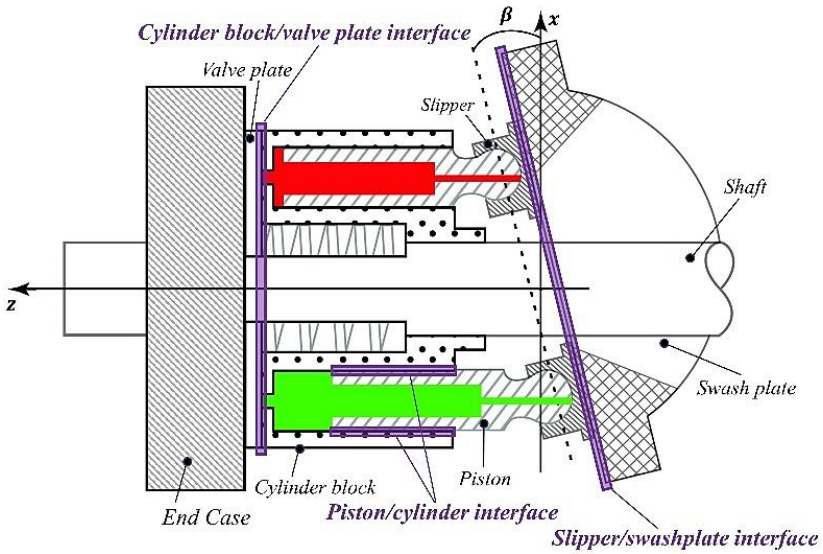


Figure 2.7 Lubricating gaps in an axial piston pump.

A reduction in gap height results in decreased viscous friction in hydrodynamic regime and increased contact friction in boundary and mixed regimes. To calculate the shear range, it was assumed that there were no-slip boundary conditions, the shear profile across the lubricating gap was linear, and pressure in the gaps was independent of gap height. In addition, it was also assumed that the fluid was incompressible and laminar flow conditions were prevalent. Hence, the shear rate in these gaps can be given as shown in Equation (2.9).

$$\frac{\partial v}{\partial h} = \lim_{h \to 0} \frac{v(h + \Delta h) - v(h)}{\Delta h} \tag{2.9}$$

Here, Δh is the constant lubricating gap height, and v is the velocity of the fluid in the gap. The pump used in this project was a 45.00 cm^3 variable displacement axial piston pump that had a pitch circle diameter of 67.00 mm and maximum swash plate angle of 18.6°. For this pump, the shear rate for the rotational speed range 600 to 2200 rpm was determined as listed in Table 2.4. Gap heights were estimated based upon the publications of Ivantysynova [135,136].

Table 2.4 Estimated range of shear rates at critical lubricating gaps.

Lubricating Gaps	Gap Height [µm]	Min. Shear Rate v_{min}/h_{max}, [1/s]	Max. Shear Rate v_{max}/h_{min}, [1/s]
Pistons/Cylinder	5.0 – 8.0	8.85×10^4	5.19×10^5
Slipper/Swash plate	7.0 – 25.0	8.42×10^4	1.10×10^6

Cylinder/Valve plate	0.9 – 2.1	1.00×10^6	8.58×10^6

Based the shear rates reported in Table 2.4, it can be assumed the approximate critical shear rate range that accounts for most of the leakage flow and viscous friction in the pump is between 8.4 x 10^4 and 8.6 x 10^6 1/s. Under idle conditions, the shear rates in the piston/cylinder and slipper/swashplate interfaces are estimated to be less than 10^5 1/s while the cylinder/valve plate shear rate approaches 10^6 1/s. Under high pressure and speed conditions, the shear rate in the piston/cylinder interface is 5 x 10^5 1/s while the shear rate in the cylinder/valve plate gap is nearly 9 x 10^6 1/s. Hence, during normal operation, shear rates within a pump can vary by as much as two orders of magnitude. If VMs can be used to tune the viscosity of the fluid at these shear rates, pump performance might be improved via the design of optimized fluids.

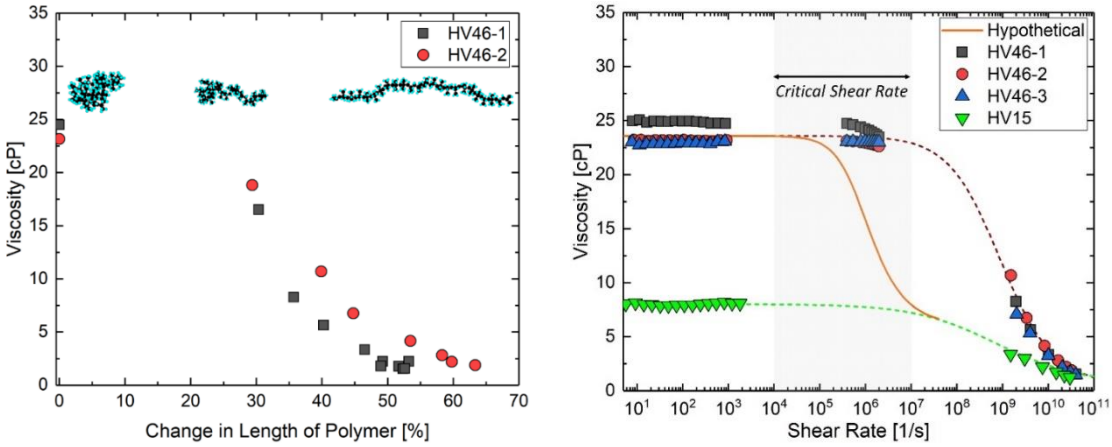


Figure 2.8 (a) Correlation between solution viscosity and the change in length of the polymer in the shear direction from the simulations. Inset are representative snapshots of the PIB as it stretches out in response to shear. (b) Viscosity as a function of shear rate from the viscometers and the MD simulations for the four fluids studies here as well as a hypothetical fluid where the VM is designed to shear thin at the critical shear rate range identified for optimum pump performance.

The proposed concept is that shear thinning exhibited by polymer-enhanced fluids might be leveraged to achieve lower viscosities in critical lubricating gaps in the pump. We showed in a previous study that temporary (as opposed to permanent) viscosity loss is important for volumetric losses [40]. Temporary viscosity loss is due to elongation of the polymers in the direction of shear such that they can resist shear force less, which causes a decrease in viscosity [43]. The simulations were used to demonstrate this effect by correlating the increase in the length of the PIB in the shear direction to the viscosity at different shear rates. As shown in Figure 2.8(a), solution viscosity was directly

correlated to polymer elongation. This has a more dramatic effect on the HV46-1 because of the higher concentration of polymer in that solution.

The simulation results show that viscosity is related to the response of the polymer to shear. This response is expected to be affected by the structure of the molecule. For example, branched polymers will be less able to elongate under shear than their linear counterparts. This suggests that a polymer might be selected or designed to respond to shear in a specific way. Specifically, this paper has shown that an ideal polymer response would be one that leads to shear thinning just after the critical shear rate range calculated for the pump. This idea is illustrated in Figure 2.8(b) where the measured viscosity for the four fluids is plotted as a function of shear rate at 50°C. The data were fitted to the Carreau-Yasuda model [137], given in Equation (2.10):

$$\mu(\dot{\gamma}) = \mu_\infty + (\mu_0 - \mu_\infty)[1 + (\lambda\dot{\gamma})^a]^{\frac{1-n}{a}} \tag{2.10}$$

Here, μ_0 and μ_∞ are the viscosity at zero and infinite shear, respectively, λ is the reciprocal of the shear rate at the onset of shear thinning, a is the Yasuda parameter that controls the transition to shear thinning, and n is the power-law exponent determining the rate of shear thinning.

None of the fluids studied here exhibited any substantial shear thinning in the critical shear rate range, but the viscosity of the HV15 was lower, corresponding to the least torque required by the pump and the most leakage flow in the pump. Therefore, the mechanical efficiency of pump was significantly higher for HV15 as compared to the HV46 fluids. However, the volumetric efficiency of the pump for HV15 is considerably lower than that for the HV46 fluids which results in the lowest overall efficiency of the pump for HV15. If a VM could be designed that enables the fluid to shear thin at the critical shear rate range, the least leakage flow and torque input could be achieved that can replicate the volumetric efficiency of the HV46 fluids and mechanical efficiency of HV15 fluid. This will result in increased overall efficiency of the pump for that fluid. The simulations, which enable explicit control of polymer chemistry and structure, as well as prediction of the elongation response of that polymer to shear, is an ideal tool for design of such tuned polymers.

2.4. Conclusion

In hydraulic systems, VMs may be subjected to high shear, causing temporary and permanent viscosity loss. VMs positively impact hydraulic efficiency by reducing internal leakage and viscous friction. The reduction in viscous friction is believed to be due to beneficial shear thinning and low traction arising from

the use of lower viscosity base oils. This work investigated the effects of diluent oil viscosity and polymer shear thinning on pump torque. Three out of the four fluids were formulated with varying base oil and polymer additive compositions such that the overall viscosities were approximately same. A considerably lower viscosity fluid consisting of a non-polymer additive composition was chosen as a fourth fluid.

The fluids were evaluated in a hydraulic dynamometer under a wide range of conditions. A statistical analysis of the mean torque and flow rates for all four fluids indicated that fluid selection did not affect pump performance when the pump was operating at 50% pressure, speed, and displacement. Differences in fluid performance were observed at operating extremes. Tests under idling conditions showed that the three HV46 fluids exhibited similar torque losses, while the HV-15 fluid required 30% less torque under idle conditions. Likewise, at high pressures, the three HV46 fluids exhibited similar torque losses, while the HV-15 fluid lower. However, the ISO VG fluid without polymer (HV46-3) exhibited slightly higher torque than the polymer containing fluids (HV46-1 and HV46-2). Empirical models for pump flow and torque were created to study the effects of diluent oil viscosity and shear thinning. An analysis of model residuals indicated that the reduction in torque produced a corresponding reduction in flow rate. Hence, the use of a low viscosity diluent oil did not impart beneficial shear thinning and instead, the overall viscosity of the fluid determined pump performance.

The critical shear rate range of the 45cc pump for the operating range was found to be 10^4—10^7 1/s. Measurements of the fluids' viscosities using rheometers at two different shear rates showed that none exhibited significant shear thinning in this range, because of the low molecular weight of the polymer. MD simulations calculated viscosity at extremely high shear rates, not accessible to experiments, which indicated that shear thinning of these fluids may occur, but at rates higher than expected in the pump. MD simulations also provided insight into the relationship between polymer conformation and temporary shear thinning. A key outcome of this study is that using a polymeric VM to enable formulation with lower viscosity base oil does not improve mechanical efficiency, at least with the low molecular weight PIB studied here. Also, the study suggested that polymers might be designed that enable a fluid to shear thin in the critical shear rate range, thereby improving mechanical efficiency without compromising the volume flow rate. Future investigations will focus on fluids formulated with higher molecular weight viscosity modifier polymers that are more susceptible to shear thinning. These findings help provide a basis for the rational formulation of polymer enhanced hydraulic fluids.

Chapter 3. Critical Shear Rate of Polymer-Enhanced Hydraulic Fluids

3.1. Introduction

The temporary shear thinning behavior presents an opportunity to tune the viscosity of hydraulic fluids. To maximize volumetric efficiency, high viscosity is required; the opposite is needed to maximize mechanical efficiency. For a fluid with VMs that exhibits shear thinning, both objectives could be achieved if the fluid is subject to low shear where flow loss is a concern and then subject to high shear where viscous friction is an issue. The challenge, however, is to (i) identify the range of shear rates relevant to hydraulic machines and (ii) formulate a hydraulic fluid that shear thins in this same range. In a previous study [79], we analyzed the operation of an axial piston pump and found the range of shear rates in key lubricating gaps within the pump to be between 10^4 and 10^7 1/s. The efficiency of the pump with three different fluids was tested, but none was found to affect efficiency because their critical shear rate was above the range calculated for the pump. However, it was suggested that if a fluid were to be formulated with higher molecular weight VMs that are more susceptible to shear thinning in the critical shear rate range, shear thinning might be leveraged to optimize the efficiency [79]. The first step towards this goal is to identify a fluid with the desired shear thinning behavior.

The key parameter needed to tune shear thinning behavior is the critical shear rate. Critical shear rate $\dot{\gamma}_{cr}$ can be correlated to the inverse of the longest relaxation time for the molecules at equilibrium (i.e., in the absence of shear), which is expected to be the rotational relaxation time λ for polymers [98,138–143]. The dynamics of polymers are typically described in terms of the Rouse [144,145] and reptation models [146–148]. The Rouse model provides an excellent description for the melt of short, unentangled polymer chains, while the dynamics of the long entangled chains can be described by the reptation model [138,147,149]. Hydraulic fluids are formulated with relatively short, unentangled polymer chains, so the Rouse model should be most applicable. In this model, the rotational relaxation time is given by $12\eta_0 M/(\pi^2 \rho R_g T)$, where η_0 is the Newtonian viscosity of the liquid, M is the molecular weight, ρ is the mass density, R_g is the universal gas constant, and T is temperature. The relationship was corroborated by molecular dynamics (MD) simulations of linear polymers with 5 to 400 monomers that showed dependence of polymer dynamics on molecular weight consistent with the Rouse model [138]. Therefore, the Rouse equation can be used to predict the approximate relaxation time of a fluid based on the molecular weight of the polymer. However, the approach has not

been extended to formulate fluids with desired relaxation times and acceptable Newtonian viscosity for hydraulic machines.

Here, we propose a model based on the Rouse equation that captures the effects of both the molecular weight and concentration of the polymer and the viscosity of the base oil to predict the formulation of a hydraulic fluid that shear thins at a predefined shear rate and with a desired Newtonian viscosity. A fluid was identified based on these predictions comprising polyisobutylene (PIB) polymers in polyalphaolefin (PAO) synthetic base oil. This fluid, along with two others with different polymer concentrations and molecular weights, were synthesized, characterized, and modeled using MD simulation. The critical shear rates predicted by the theoretical model were compared to rheological data across several decades of shear rates obtained from experiments and simulations. MD simulations were then used to explore the molecular origins of the shear thinning behavior. This study lays the groundwork for design of fluids with tunable viscosity behavior and optimized hydraulic efficiency.

3.2. Materials and Methods

3.2.1. Rheological Measurements

Two different instruments (see Figure A.1) were used to characterize the rheological properties of the test fluids at 50 and 80 °C in both the Newtonian and non-Newtonian regimes. First, a Cannon StressTech HR Oscillatory Rheometer was used to measure the dynamic viscosity of the fluids under low shear rate conditions. The shear rate was varied from 10 to 1000 1/s, with 15 measurements taken at logarithmically spaced intervals in that range. This procedure was repeated three times for each fluid to report average viscosity values at each shear rate and temperature. The standard deviation of the average of the three measurements was around ±0.2 mPa·s.

A PCS Ultra Shear viscometer was used to measure the high-shear viscosity at approximate shear rates from 300,000 to 4,000,000 1/s. Two repeat high-shear viscosity measurements were performed for one of the fluids and the difference was used to estimate a maximum error of 1.4 mPa·s. Measurements at the highest shear rates also exhibited a temperature rise up to 5 °C above the set temperature (50 or 80 °C). This may have resulted in artificially low viscosity measurements at the highest shear rates.

3.2.2. Molecular Dynamics Simulation

MD simulations were used to obtain the Newtonian viscosity, high-shear viscosity, and to understand the dynamics of the molecules during shear. The model systems were created using Material Studio software. All atomic interactions were described using a united atom (UA) potential. UA potentials are known to accurately describe the viscosity of hydrocarbons at ambient conditions and high pressures according to numerous past studies [99,124,139,140,150–152]. The UA parameters were selected from Transferable Potentials for Phase Equilibria (TraPPE-UA) developed for branched alkanes and alkenes [153–156]. Details about the UA potential and force field parameters are given in Section A.2. Simulations using the TraPPE-UA force field were performed with a spherical cutoff of 1.4 nm and analytic tail corrections for the Lennard–Jones interactions. The bond length is not considered to be rigid but is controlled by a harmonic potential with the force constants taken from the CHARMM force field [156–158]. For all simulations, a time step of 1.0 fs was used, and periodic boundary conditions were applied in all directions with a 4 × 4 × 30 nm simulation box. The box was longer in one direction to allow the molecules to elongate in response to shear in high-shear viscosity simulations without unphysical effects from the periodic boundary. Dynamic simulations were run using Large Atomic/Molecular Massively Parallel Simulation (LAMMPS) software [121].

To prepare the model system, an energy minimization of the system was performed using the conjugate gradient algorithm. The system density was then equilibrated at 1.0 atm and either 50 or 80 °C for 50 ns in the isothermal–isobaric (NPT) ensemble using the Nosé–Hoover thermostat and barostat [125,126], with damping coefficients of 100 and 125 fs, respectively. Then, while maintaining a constant temperature for 10 ns in the canonical (NVT) ensemble, the simulation box was deformed until the density of the fluid reached the average density computed from the last 10 ns of the previous NPT simulations. Finally, the system was equilibrated using the final configuration from NVT as the initial configuration for 5 ns in the microcanonical (NVE) ensemble. This way, a desired equilibration state of the system was achieved while avoiding interference with the dynamics of the system [159]. Following this equilibration process, two sets of simulations, equilibrium, and non-equilibrium, were carried out to calculate viscosity at 50 and 80 °C. These temperatures were chosen to reflect the temperature range experienced by lubricants used as hydraulic fluids.

First, the Newtonian viscosity of fluids was modeled using equilibrium molecular dynamics (EMD). Specifically, we used the Green–Kubo (GK) approach, which relates the viscosity to the time integral of the stress

autocorrelation function [159,160]; see Section A.3 for details. Using the previously equilibrated model, simulations were equilibrated again for 2 ns in the NPT ensemble followed by NVE ensemble production runs. The pressure tensor components were saved every 5 fs for 5 ns at 50 °C and 20 ns at 80 °C during the NVE simulation. An average low-shear viscosity value was calculated from multiple NVE trajectories and by applying the GK formalism following the time decomposition approach [160].

Second, non-equilibrium molecular dynamics (NEMD) simulations were used to determine the viscosity at high shear rates and to study the non-Newtonian viscosity behavior. NEMD simulations for calculating viscosity involve applying a planar Couette flow field at shear rate $\dot{\gamma} = \partial u_x/\partial z$ to induce a flow gradient or momentum flux, and characterizing the linear change in streaming velocity of flow u_x in the x-direction with vertical position z [11,129,130]. First, shear was imposed on the simulation box at a desired strain rate by deforming the box in the x-direction for 100–1000 ns, depending on the shear rate. At the same time, the fluids were thermostatted via a Nosé Hover thermostat and the SLLOD equations of motion. Depending on the shear rate, 10–200 ns of time was provided for the system to achieve steady-state before the shear stress P_{xz} and velocities of the atoms in the direction of shear u_x were collected for post processing. It was assumed that an NEMD simulation reached steady state if the goodness of the slope of a linear fit to the average atom velocities was 99% or greater. The average shear rate was then calculated from the slope of a linear fit to the average atom velocities and average viscosity was determined from the ratio of the average shear stress to the average shear strain rate; see Section A.3 for details. This process was repeated for each test fluid at 50 and 80 °C for shear rates ranging from 1×10^8 to 1×10^{11} 1/s.

3.3. Results and Discussion

3.3.1. Fluid Identification and Formulation

We considered blends of linear polyisobutylene (PIB) polymer and synthetic polyalphaolefin (PAO) base oil. The Newtonian viscosities and critical shear rates were calculated for fluids comprising 12, 16, or 20 wt.% PIB polymer at varying concentrations in PAO 2 mm²/s (PAO2), PAO 4 mm²/s (PAO4) or PAO 8 mm²/s (PAO8) base oil. The PAOs used here are representative of the viscosity grades of base stocks used in formulating hydraulic fluids that usually range from 2–8 mm²/s at 100 °C, depending on the additives.

The Newtonian viscosity η_0 and density ρ of the blend were predicted by the Kendall–Monroe relation [161] and the sum of the mass fractional density of

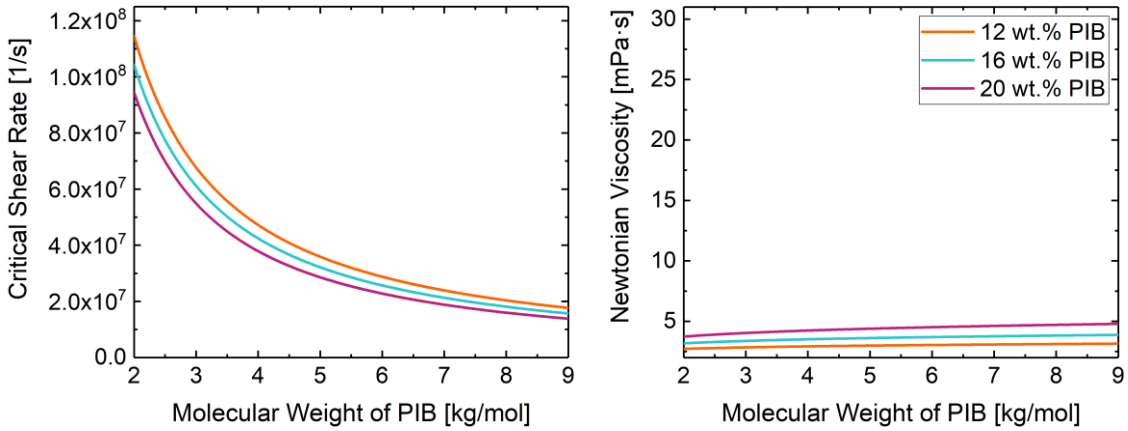
each component, respectively, using the viscosities (η_{0p} and η_{0s}), densities (ρ_p and ρ_s), and concentrations (c_p and c_s) of the polymer p and base oil solvent s. The Newtonian viscosity η_{0p} of the polymers at 25 °C as a function of molecular weight was calculated from an empirical model derived from a large set of experimental data [162]. The viscosity of the polymers at 25 °C was then used to obtain viscosity at 80 °C using the temperature shift factor for PIB [163] and the temperature dependence of viscosity described by the William–Landel–Ferry equation [163]; all equations are given in Section A.4.

The critical shear rate was calculated using a model developed by combining the Rouse equation [140,145,164] for rotational relaxation time with the kinetic-theory-based rigid dumbbells model proposed by Bird et al. [11,165], which is similar to the Rouse model but includes polymer concentration and base oil viscosity terms; original equations given in Section A.4. The proposed model correlates critical shear rate $\dot{\gamma}_{cr}$ to the molecular weight M and concentration c_p of the polymer, and the viscosity and density of the blend and base oil as

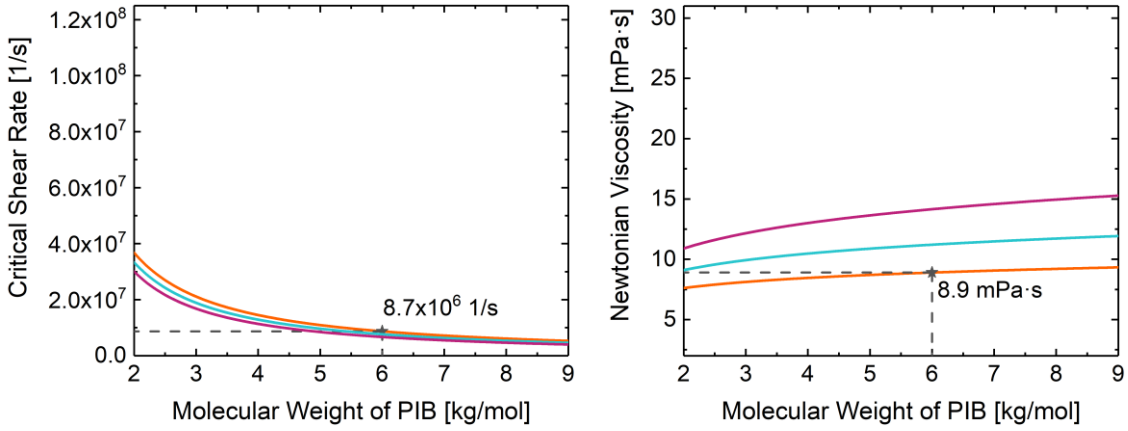
$$\dot{\gamma}_{cr} = 1/\lambda = \frac{\pi^2}{12}\frac{c_p \rho R_g T}{(\eta_0 - \eta_{0s})M} \tag{3.1}$$

Here, R_g is the universal gas constant and T is the temperature. These models were originally derived based on the assumption that only a single molecular mass, in this case the polymer, contributed to the shear response of the blend. The same is true for this new model. For a liquid, shear thinning is said to occur when the Weissenberg number Wi, which is the product of the relaxation time λ and shear rate $\dot{\gamma}_{cr}$, is greater than unity (that is, $Wi = \lambda\dot{\gamma}_{cr} > 1$) [11]. This criterion was used with Equation (3.1) to estimate the critical shear rate for different combinations of commercially available PAO base oils and PIB polymers across a range of polymer concentrations.

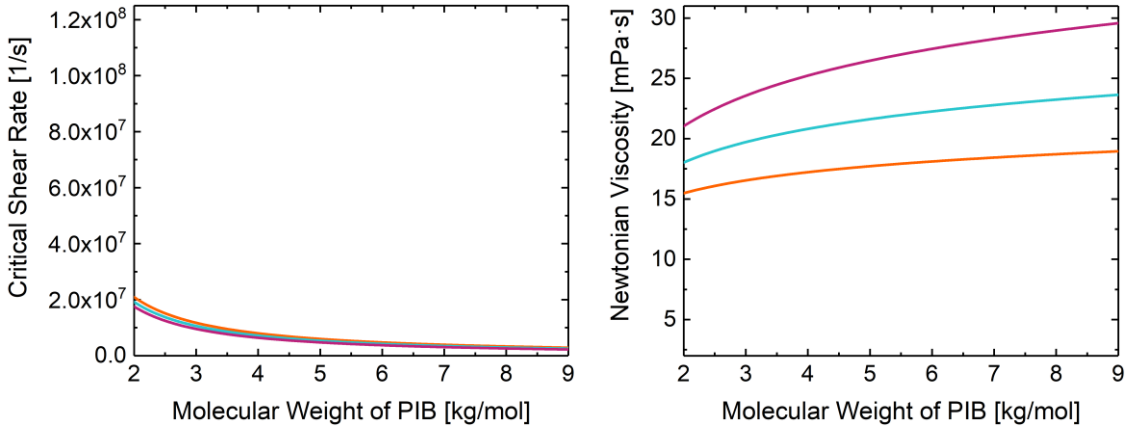
Figure 3.1 shows the calculated critical shear rate and Newtonian viscosity of fluids comprising 12, 16, and 20 wt.% PIB polymer with PAO2, PAO4, or PAO8 base oil as a function of the molecular weight of the polymer. The critical shear rate decreases with both the molecular weight and concentration. However, the effect of polymer concentration on the critical shear rate is small relative to the effect of molecular weight. The Newtonian viscosity increases with molecular weight and concentration. Unlike critical shear rate, the effect of molecular weight is more significant than that of concentration. It can also be observed that a fluid blended with a lower viscosity base oil has a higher critical shear rate and lower Newtonian viscosity fluid.



(a)



(b)



(c)

Figure 3.1 Predicted critical shear rate and Newtonian viscosity at 80 °C of polymer and base oil blends as functions of polymer molecular weight for three different polymer concentrations in (a) PAO2, (b) PAO4, and (c) PAO8. The ideal fluid is highlighted by the dashed lines and stars in (b).

The predictions in Figure 3.1 can be used to identify an ideal fluid formulation. Specifically, the target blend should have a Newtonian viscosity suitable for a hydraulic pump (8 to 10 mPa s at 80 °C) and critical shear rate in the range of shear rates exhibited in the key lubricating gaps within that pump (10^4 to 10^7 1/s). The model predictions suggest that an ideal fluid formulation is 12 wt.% 6.0 kg/mol PIB in PAO4, identified by dashed lines and the star in Figure 3.1(b). This blend should have a critical shear rate of 8.7×10^6 1/s and Newtonian viscosity of 8.9 mPa s at 80 °C, both of which are within the desired range for the hydraulic pump.

Based on the outcomes of this analysis, we formulated a fluid blend (Fluid 1) from 12 wt.% 6.0 kg/mol PIB (PIB6000) polymer with 88 wt.% PAO4 base oil. Then, two additional fluids were designed to test the effects of concentration and molecular weight of the polymer on critical shear rate. Fluid 2 was formulated with 16 wt.% PIB6000 in PAO4 while Fluid 3 was formulated with 12 wt.% of 1.3 kg/mol polyisobutylene (PIB1300) polymer in PAO8 base oil. A summary of the formulation plan for these three fluids is given in Table 3.1. Fluids 1 and 2 have the same polymer and base oil, but different polymer concentration. Fluids 1 and 3 have the same polymer concentration, with different molecular weight polymer and PAO base oil. Apart from the polymer, no other additives were included in the fluid formulations.

Table 3.1 Chemical composition of the test fluids.

Fluid ID	Fluid 1	Fluid 2	Fluid 3
PAO4	88 wt.%	84 wt.%	-
PAO8	-	-	88 wt.%
PIB1300	-	-	12 wt.%
PIB6000	12 wt.%	16 wt.%	-

The three fluids formulated above were modeled using MD simulations. Consistent with previous gas chromatography-mass spectroscopy measurements, PAO4 was modeled as equal parts 1-decene trimer and tetramer and PAO8 was modeled as 1-decene tetramer [119]. The PIB1300 polymer was built from 22 isobutylene monomers and an exo group so that the resulting molecule had a nominal molecular weight of 1300 g/mol and a bromine number of 12, indicating it was a linear olefin with one unsaturated bond ($C_{96}H_{192}$) [120]. Similarly, the PIB6000 polymer comprised 106 isobutylene monomers and an exo group so that the resulting molecule had a nominal molecular weight of 6000 g/mol and a bromine number of 3, indicating it was a linear olefin with one unsaturated bond ($C_{428}H_{856}$) [120]. The virtual models of the 1-decene molecules and PIB polymers are shown in Figure 3.2.

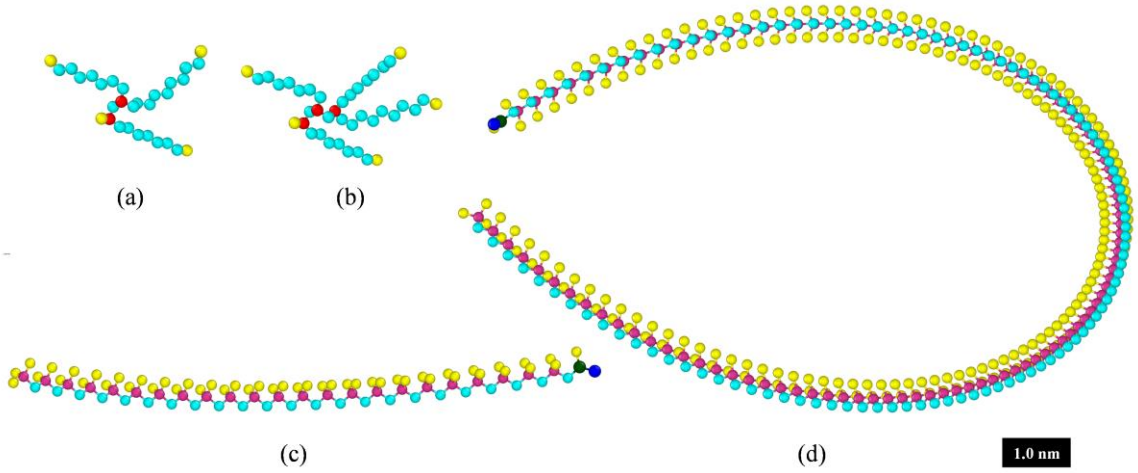


Figure 3.2 Molecular structures of (a) decene trimer, (b) decene tetramer, (c) PIB1300, and (d) PIB6000 in united atom representation. The yellow, cyan, red, pink, blue, and green spheres represent CH3, CH2, CH, C, sp2 CH2, and sp2 C, respectively.

The three model fluid systems were created by placing base oil and polymer molecules in a periodic simulation box. The number of molecules was determined based on the percent weight concentrations in Table 3.1. Note that the number of molecules in each model system varied, but the size of the simulation box was kept approximately constant. A summary of the number of molecules in the model of each of the three fluids is reported in Table 3.2.

Table 3.2 Simulation plan for the three fluids with the number of molecules of each type for each model.

Fluid ID	Fluid 1	Fluid 2	Fluid 3
Decene trimer, $C_{30}H_{62}$	261	244	-
Decene tetramer, $C_{40}H_{82}$	196	168	386
Polymer, PIB1300 ($C_{96}H_{192}$)	-	-	22
Polymer, PIB6000 ($C_{428}H_{856}$)	5	6	-

3.3.2. Rheological Behavior

The accuracy of the simulations was evaluated by comparing Newtonian viscosities obtained from the EMD simulations to measured values for the fluids. The Newtonian viscosities of these fluids as a function of simulation time at 50 and 80 °C with their standard deviations are shown in Figure 3.3. The steady-state viscosity of the fluids at 50 and 80 °C was calculated from the average of the running integral of the Green–Kubo formula over 40 and 20 NVE

trajectories, respectively. The viscosities from the simulations are compared to measurements from the Cannon StressTech HR Oscillatory Rheometer in Table 3.3. This comparison shows that the simulation and measured viscosities are the same within the simulation and experimental error for both fluids at both temperatures. This indicates that the model systems are good representations of the formulated fluids.

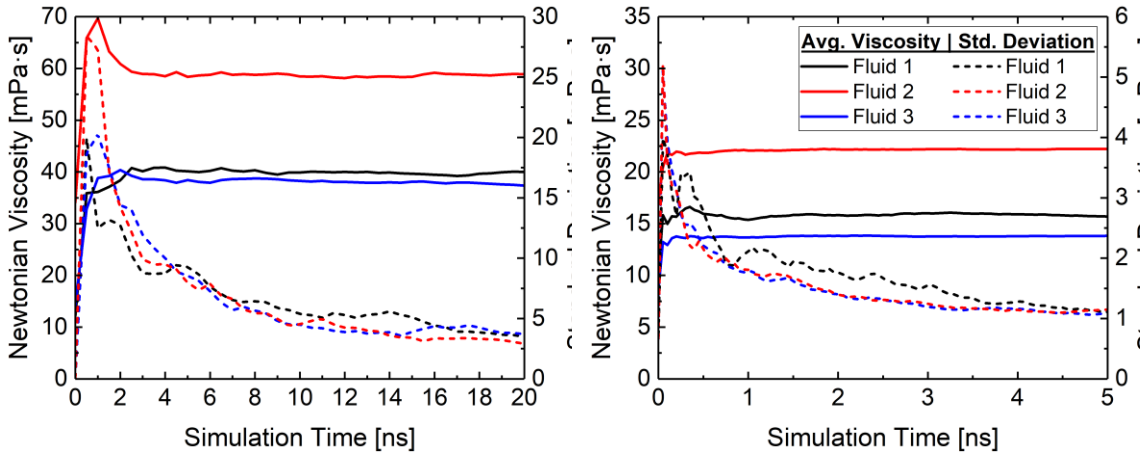


Figure 3.3 Average Newtonian viscosity of the three fluids at (a) 50 °C and (b) 80 °C and standard deviation obtained from 20 NVE trajectories in equilibrium molecular dynamics (EMD) simulations.

Table 3.3 Comparison of simulated Newtonian viscosity of the three fluids with experimentally Newtonian viscosities. The error associated with all experimental data is ±0.2 mPa·s.

Temperature	Fluid ID	Viscosity (mPa·s)	
		Experiment	Simulation
50 °C	Fluid 1	40.2	39.7 ± 4.0
	Fluid 2	62.3	58.9 ± 3.2
	Fluid 3	35.5	38.6 ± 3.4
80 °C	Fluid 1	15.9	15.7 ± 1.1
	Fluid 2	22.3	23.1 ± 1.0
	Fluid 3	13.5	14.3 ± 1.0

Next, we characterized the high-shear viscosity for the fluids using the PCS Ultra Shear viscometer at shear rates from 3×10^5 to 4×10^6 1/s and NEMD simulations from 1×10^8 to 1×10^{11} 1/s. These data were then combined with the Newtonian viscosity from EMD simulations (approximated as 1×10^0 1/s) and the Cannon rheometer at shear rates of 7×10^0 to 1×10^3 1/s to generate a complete viscosity profile across a wide range of shear rates. The results are shown in

Figure 3.4, where hollow and solid symbols represent data from experiment and simulation, respectively. At both 50 and 80 °C, all three fluids exhibit constant Newtonian viscosity at low shear rates, and then shear thin at high shear rates, as expected. The viscosity as a function of shear rate data $\eta(\dot{\gamma})$ was then fit to the Carreau equation for shear thinning [166]:

$$\eta(\dot{\gamma}) = \eta_\infty + (\eta_0 - \eta_\infty)[1 + (\lambda\dot{\gamma})^2]^{\frac{n-1}{2}} \tag{3.2}$$

Here, η_∞ is the viscosity at infinite shear rate and n is the power-law exponent. All viscosity data from simulations and experiments was fit to Equation (3.2) as shown by the dashed lines in Figure 3.4. The critical shear rates, the inverse of the fitted relaxation times, are reported in Table 3.4. The complete setof fit parameters for Equation (3.2) is in Table A.5.

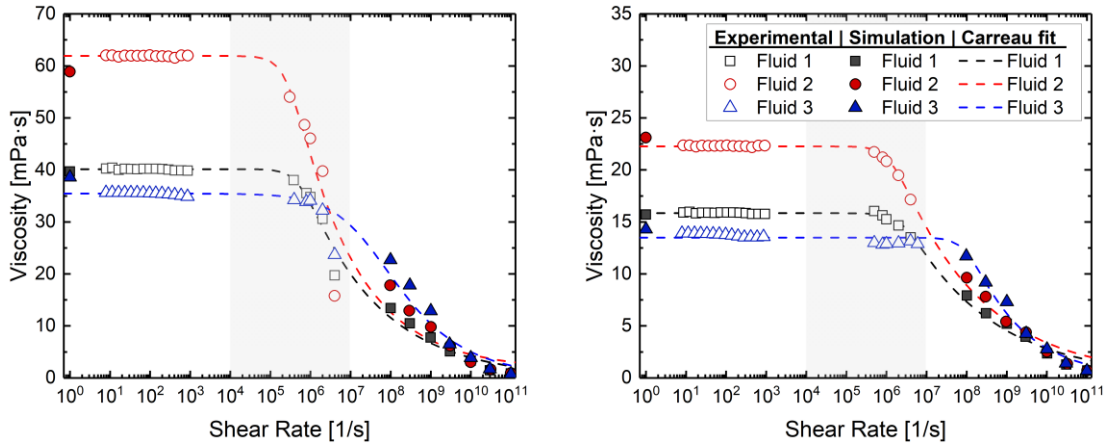


Figure 3.4 Viscosity as a function of shear rate from experiments at two different shear rate ranges (hollow symbols) and simulations (solid symbols) for the fluids studied at (a) 50 °C and (b) 80 °C. The dotted lines are the Carreau model fit to the viscosity profile of each fluid. The shaded region indicates the range of shear rates in the key lubricating gaps within the pump.

The onset of shear thinning for Fluids 1 and 2 occurs in the key shear rate range for hydraulic pumps, indicated by the shaded regions in Figure 3.4. These two fluids were formulated with the same molecular weight PIB but different polymer concentrations and have approximately the same critical shear rate at 80 °C. However, Fluid 2 has a lower critical shear rate at 50 °C. In addition, Fluid 2 has significantly higher Newtonian viscosity both at 50 and 80 °C than Fluid 1 due the larger concentration of polymer. We also observe that the critical shear rate of the fluids at 50 °C is lower than that at 80 °C. This is due to dependence of relaxation time and viscosity on temperature (Equation (3.1)). These results confirm that molecular weight governs critical shear rate, whereas concentration governs the Newtonian viscosity and the rate of shear

thinning. This behavior is consistent with the predictions from the theoretical analysis.

The critical shear rates from the Carreau fit to experiment and/or simulation data for fluids arecompared to the predictions from Equation (3.1) in Table 3.4. This comparison shows good agreement between prediction and observation. The minor differences are attributable to the assumption made in deriving Equation (3.1) that a single molecular weight contributes to the shear response of thefluid, which neglects the contribution of base oil molecules. The model also does not account for the architecture of molecules that contributes to the shear response. However, the results indicate that wewere successful in using a simple model to identify a fluid formulation that shear thins in a specific range of shear rates, specifically the range of shear rates relevant to hydraulic machines.

Table 3.4 Critical shear rate of the fluids obtained from the Carreau fit to experiment and simulation viscosity data as compared to the theoretical prediction.

Temperature	Fluid ID	Critical Shear Rate (1/s)	
		Experiment/ Simulation	Prediction
50 °C	Fluid 1	5.27×10^6	1.19×10^6
	Fluid 2	62.3	58.9 ± 3.2
	Fluid 3	35.5	38.6 ± 3.4
80 °C	Fluid 1	15.9	15.7 ± 1.1
	Fluid 2	22.3	23.1 ± 1.0
	Fluid 3	13.5	14.3 ± 1.0

To understand the observed shear thinning behavior, the simulations were used to analyze the change in molecular conformation due to shear stress. Temporary viscosity loss is expected to be caused by elongation of molecules in the direction of shear such that the molecules can provide less shear resistance, i.e., lower viscosity. To quantify this, we calculated the change in polymer length in the direction of shear from the NEMD simulations at each shear rate, averaged over time for all polymers in the simulation. The results are plotted vs. normalized viscosity (shear viscosity divided by Newtonian viscosity from EMD simulations) in Figure 3.5. Consistent with the expected mechanism of temporary viscosity loss, greater change in molecule length corresponds to lower viscosity.

Comparing the different fluids in Figure 3.5, at both temperatures, the average change in polymerlength is much larger for Fluids 1 and 2 than for Fluid 3. This can be explained by the fact that Fluids 1and 2 contain higher molecular weight polymers. These longer polymers require more time to respond to a given shear

stress, i.e., they have longer relaxation times, and therefore have a lower critical shear rate (Table 3.4). The polymer length results also show that elongation depends on temperature. Particularly for Fluids 1 and 2, the change in polymer length is greater at 80 °C than 50 °C. This can be explained by the fact that there is more kinetic energy at a higher temperature, which enables the molecules to respond faster to shear. A faster response to shear corresponds to a shorter relaxation time and therefore a higher critical shear rate at 80 °C.

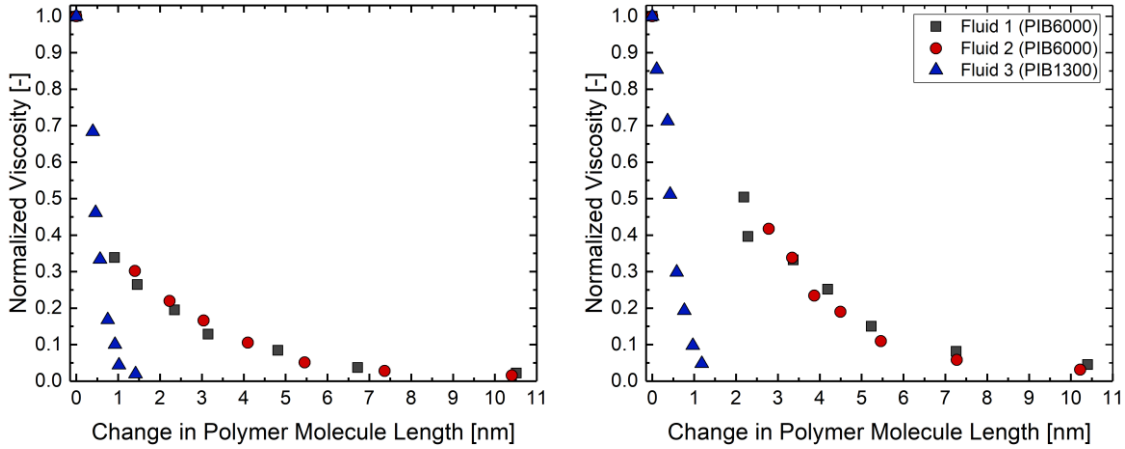


Figure 3.5 Normalized viscosity as a function of change in average polymer length in the direction of shear rate for the fluids studies here at (a) 50 °C and (b) 80 °C. Each point corresponds to a different shear rate.

3.4. Conclusion

Energy-efficient commercial hydraulic fluids are formulated with polymeric VMs that cause the fluids to exhibit shear thinning. While temporary viscosity loss is often viewed as detrimental, here it was explored as a potentially beneficial means of tuning the viscous behavior of a fluid. Particularly, we considered the possibility of using shear thinning to optimize the efficiency of hydraulic pumps for which efficiency is a non-monotonic function of viscosity. The ideal fluid for a hydraulic pump was identified as one with a critical shear rate between 10^4 and 10^7 1/s and the Newtonian viscosity of 8–10 mPa·s at 80 °C. A theoretical model was developed to predict critical shear rate and Newtonian viscosity as a function of the properties of the VM polymer and base oil, and their relative concentrations. The model was applied to PIB polymers of varying molecular weights in several different PAO base oils. This model predicted that a fluid formulated using 12 wt.% of 6 kg/mol PIB with PAO4 base oil should have a critical shear rate lower than 10^7 1/s with Newtonian viscosity of 8.9 mPa·s.

Based on the theoretical predictions, three fluids were formulated to validate the model and to distinguish the effects of molecular weight and concentration of polymer on the critical shear rate of the fluids. Fluids 1 and Fluid 2 were formulated using PAO4 base oil and 12 wt.% and 16 wt.% of 6.0 kg/mol PIB, respectively, and Fluid 3 was formulated using PAO8 base oil and 12 wt.% 1.3 kg/mol PIB. The viscosities of all three fluids across a wide range of shear rates were obtained using experimental measurements and atomistic simulations at 50 and 80 °C. The low shear viscosities calculated from the simulations were consistent with those measured experimentally. The results indicate that the molecular weight of the polymer plays a key role in determining the critical shear rate, whereas the concentration of the polymer dominates the Newtonian viscosity. It was also found that the critical shear rates predicted by the theoretical model were consistent with the calculated critical shear rate using measured and simulated viscosities of fluids. This indicates that this model can be used more generally to predict critical shear rates of fluids and also formulate fluids that exhibit desired shear thinning profiles.

Finally, the shear thinning behavior of fluids was studied using the MD simulations. The simulations showed a direct correlation between the average change in polymer molecule length in the direction of shear to the normalized viscosity of fluids. The findings also confirmed that fluids formulated using longer polymers had lower critical shear rates because of their longer relaxation time, i.e., they take longer to respond to shear. Further, the simulations explained that critical shear rate is larger at higher temperatures because the additional kinetic energy enables fluids to respond to elongate in response to shear more quickly.

Overall, the results of this study and findings from our previous dynamometer study [79] suggest that Fluid 1 might increase mechanical efficiency without compromising the volumetric efficiency of the hydraulic pumps, since it is adequately susceptible to shear thinning in the key shear rate range of a pump and has an acceptable Newtonian viscosity for hydraulic machines. This is not the case for Fluid 2 because of its high Newtonian viscosity, nor for Fluid 3 that shear thins at a shear rate above that expected to occur in the pump. To validate this prediction, pump performance testing can be conducted in a dynamometer for these three fluids. In addition, the temperature dependency of the critical shear rate suggests that the fluid should be formulated to shear at a specific temperature range. More generally, this investigation provides a simple model for predicting critical shear rate that can be used to formulate polymer-enhanced hydraulic fluids that exhibit a shear thinning profile with a desired critical shear rate.

Chapter 4. Effect of Polymer Structure and Chemistry on Viscosity Index, Thickening Efficiency, and Traction Coefficient of Lubricants

4.1. Introduction

Many studies listed in Chapter 1 demonstrate that there is a wide range of functionality and behavior of polymers in lubricant formulations. The structure and chemistry of the polymers vary depending on their purpose and affect the mechanisms by which they function. However, the correlations between structure/chemistry and lubricant function are still poorly understood. In this context, the goal of this project was to characterize the effect of polymer chemistry/structure on VI, TE, and TC. To achieve this goal, we used MD simulations of five commercially relevant polymer-base oil formulations. The simulations were used to obtain viscosity using standard methods at 40 and 100 °C, and VI and TE were calculated correspondingly. Then, a new method was developed for simulating TC. All simulation results were validated by direct comparison to experimental measurements. Next, the polymer structure data available from the simulations was used to develop simple empirical models to predict VI, TE, and TC. Although these models require information from the simulation, they are far less computationally expensive and so can be used to predict performance metrics rapidly and with reasonable accuracy. Finally, the efficacy of the empirical models was evaluated for a sixth fluid that was not part of the model development.

4.2. Materials and Methods

4.2.1. Test Fluids

The five ISO VG32 fluids evaluated in this study are listed in Table 4.1. These fluids were formulated by blending five polymers with varying chemistries and degrees of branching with 3 mm^2/s group III base oil such that the formulated fluids had a kinematic viscosity of 7 mm^2/s at 100 °C. For this study, we chose five different branched and linear polymers with number average molar mass ranging from 3.0 to 8.5 kg/mol and viscosity ranging from 480 to 997 mm^2/s. The names of the fluids start with FKV to indicate they had similar/fixed kinematic viscosity and end with the abbreviation of the polymer additive with which they were formulated.

FKV-PAO was formulated by blending 3 mm^2/s group III base oil with 15.7 wt.% of polyalphaolefin (PAO) polymer having number average molar mass of

5.0 kg/mol, polydispersity index (PDI) of 1.1, density of 0.86 g/cc, and viscosity of 500 mm²/s. The PAO in the FKV-PAO fluid comprised 100 wt.% 1-decene monomer. The polyisobutylene (PIB) in FKV-PIB fluid had 100 wt.% isobutylene monomer. The butadiene isoprene (BDIP) polymer in FKV-BDIP fluid had 50 wt.% butadiene and 50 wt.% isoprene monomers. The polyalkyl methacrylate (PAMA) polymer in FKV-PAMA fluid had 22 wt.% C12-methacrylate, 30 wt.% C13-methacrylate, 30 wt.% C14-methacrylate, and 18 wt.% C15-methacrylate monomers of varying branching patterned of alkyl chains. The methacrylate butadiene (MABD) polymer in FKV-MABD fluid had 6.0 wt.% methyl methacrylate, 8.5 wt.% butyl methacrylate, 75.5 wt.% lauryl methacrylate, and 10 wt.% butadiene. The BDIP and MABD polymers were saturated by hydrogenation. It is notable that the PIB and the BDIP are linear whereas the remaining three polymers PAO, PAMA, and MABD have varying degrees of branching. All the fluids were formulated with 6.5 wt.% Anglamol 99 (32% S, 1.7% P, 0.05% N) antiwear and high-pressure additive package.

Table 4.1 Description of the test fluids and polymers. The colors scheme introduced here will be used to identify the fluids subsequently in this paper. Here, PDI = polydispersity index, M_n = number average molar mass, ρ = density, and ν = kinematic viscosity.

Properties	FKV-PAO	FKV-PIB	FKV-BDIP	FKV-PAMA	FKV-MABD
Base stock ID	GIII 3 mm²/s	GIII 3 mm²/s	GIII 3 mm²/s	GIII 3 mm²/s	GIII 3 mm²/s
Polymer ID	PAO	PIB	BDIP	PAMA	MABD
Monomer(s)	1-decene	isobutylene	butadiene, isoprene	C12-C15 methacrylate	methacrylate, butadiene
Treat rate, wt.%	15.7	14.6	9.8	20.8	15.8
M_n, kg/mol	5.0	2.1	3.0	8.5	4.3
PDI	1.1	1.8	2.8	1.7	2.8
ρ at 15.6 °C, g/cc	0.86	0.91	0.86	0.94	0.95
ν at 100 °C, mm²/s	500	2500	997	480	-

4.2.2. Rheological and Traction Measurements

Three different instruments (see Figure B.1) were used to measure the Newtonian viscosities and TC of the fluids. The Newtonian viscosity was measured at 40 and 100 °C so that the VI and TE of the fluids could be characterized. First, a viscometer was used to measure the kinematic viscosity of all fluids. These measurements were conducted per the ASTM D-445 test standard [115]. Second, a Cannon StressTech HR Oscillatory Rheometer was used to measure the dynamic viscosity of all fluids, except for FKV-PAO and FKV-PIB. The shear rate was varied from 10 to 1000 1/s, with 15 measurements taken at logarithmically spaced intervals in that range. This procedure was repeated three times for each fluid to report the average Newtonian viscosity values at each temperature. The standard deviation of the

average of the three measurements was around 0.2-0.3 mPa.s. Third, a PCS Instruments Mini-Traction Machine (MTM) was used to measure the traction coefficient of all fluids at 40 °C in the full film lubrication regime. The tests were performed at an entrainment speed of 1.0 m/s, load of 75 N, and slide-to-roll ratio of 20.0%. Both the ball and disk specimens were AISI 52100 steel with an elastic modulus of 207 GPa and Poisson's ratio of 0.3. The diameter of ball was 19.05 mm. The average roughness of ball and disk were 12.0 and 6.0 nm, respectively. The traction coefficient measurement of each fluid was performed once per fluid. The shear rate experienced by the fluid during these tests was estimated based on the experimental parameters and estimated film thickness, as described in the supplementary information Section B.10, to be 1.6×10^6 1/s. All the experimental data are tabulated in Section B.2 of the supplementary material. For some fluids, viscosity and traction coefficient were measured with and without the Anglamol 99 additive. The comparison of the properties of three fluids with and without 6.5 wt.% of Anglamol additive package is shown in Figures B.2 and B.3. It was found that the fluids formulated with the additive package had higher viscosity, thickening efficiency, and traction coefficient than the fluids formulated without the additive package. However, the comparison also shows that the fluids with and without additive exhibited consistent trends.

4.2.3. Molecular Dynamics Simulations

The model systems were created to reproduce the physical fluids studied experimentally (Table 4.1). The model structures of the base oil and polymer molecules are shown in Figure 4.1. The differences between the physical and model fluids were the PDI, the base oil chemistry, and the presence of the additive package. First, for the virtual formulations, a PDI of 1.0 was assumed whereas the PDI of the physical fluids was between 1.1 and 2.8. Generally in polymer melts, as the PDI increases, the Newtonian viscosity decreases and the shear thinning behavior becomes more pronounced [167]. However, in our case, this effect should be negligible due to the small concentration of the polymer and low viscosity of the solutions. Second, since the composition of group III (3 mm^2/s) base oil was unknown, a model 3 mm^2/s base oil was created from 25 wt.% PAO 2 mm^2/s (PAO2) and 75 wt.% PAO 4 mm^2/s (PAO4). Previous gas chromatography-mass spectroscopy measurements reported that PAO2 comprised nearly 100 wt.% dimer of 1-decene whereas PAO4 comprised 50 wt.% trimer of 1-decene + 50 wt.% tetramer of 1-decene [119]. More details, including viscosity measurements (Table B.3) and a sample calculation of the ratios of each molecule (Table B.4), about the formulation of the 3 mm^2/s PAO base oil can be found in Section B.3 of the supplementary material. Group III base oils usually have lower VI than group IV oils (i.e., PAO base oils) [168], so

the model fluids may have higher VI than their viscosity counterparts. Lastly, the models did not contain the additive package. Therefore, these model fluids were created with 6.5 wt.% more 3 mm^2/s group III base oil than the physical fluids. For example, the model FKV-PAO comprised 15.7 wt.% of PAO with 84.3 wt.% of 3 mm^2/s group III base oil in the FKV-PAO fluid.

The model systems were created using Material Studio software. More details about each model system, including composition, simulation box size, and number of each type of molecule, are given in Table B.5. For all simulations, a time step of 1.0 fs was used, and periodic boundary conditions were applied in all directions. An orthogonal simulation box ($30 \times 3 \times 3$ nm^3) was used to allow the molecules to elongate in response to shear applied in the direction of the long side of the box in the traction simulations. This model design minimized finite-size effects [169] where molecules might unphysically interact with themselves across the periodic boundary. Not all properties are affected by the system size; for example, the shear viscosity showed no significant system-size dependences, unlike the diffusion coefficient [170]. Nevertheless, a large enough simulation box was used to reduce pressure and stress fluctuations [171], enabling accurate and reliable calculation of the mechanical properties of a polymer system [172].

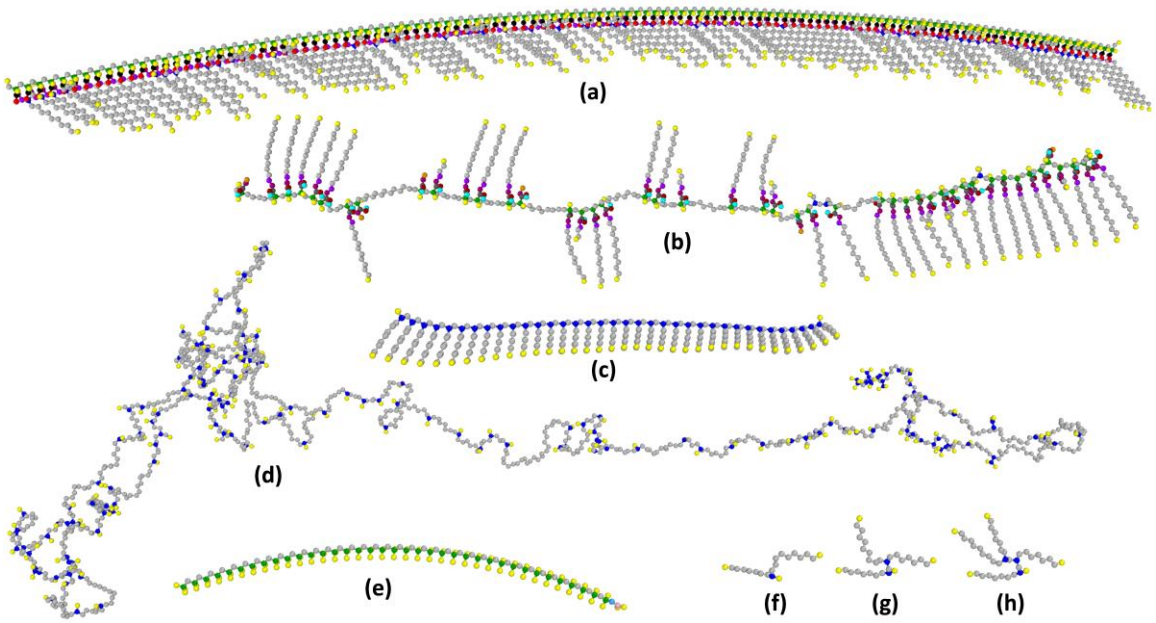


Figure 4.1 Structure of the molecules in united atom representation: (a) PAMA polymer, (b) MABD polymer (c) PAO polymer, (d) BDIP polymer, (e) PIB polymer, (f) dimer of 1-decene, (g) trimer of 1-decene, and (h) tetramer of 1-decene. The sphere colors represent the following pseudoatoms: yellow-CH_3, silver-CH_2, blue-CH, green-C, black-C(=O), red-O(=C), magenta-O, purple-CH_2(-O), orange-CH_2(-O), dark grey-CH(-OH), cyan-H(-O) sky blue-CH(=C), and pink-C(=CH).

All atomic interactions were described using a united atom (UA) potential. UA potentials are known to accurately describe the viscosity of hydrocarbons for broad range of temperatures, pressures, and shear stresses according to numerous past studies [78,99,124,139,140,151,152,173,174]. The UA parameters were selected from Transferable Potentials for Phase Equilibria (TraPPE-UA) developed for branched alkanes, alkenes, alcohol, aldehyde, ketone, methacrylate, and carboxylic acids [153,154,181,182,155,156,175–180]. Details about the UA potential and force field parameters are given in Section B.5. Simulations using the TraPPE-UA force field were performed with a spherical cutoff of 1.4 nm and analytic tail corrections for the Lennard–Jones interactions. The covalent bonds were modeled by a harmonic potential with the force constants [156,180]. Dynamic simulations were run using Large Atomic/Molecular Massively Parallel Simulation (LAMMPS) software [121].

To prepare the model system, an energy minimization of the system was performed using the conjugate gradient algorithm. The system density was then equilibrated at 1.0 atm and either 40 or 100 °C for 50 ns in the isothermal–isobaric (NPT) ensemble using the Nosé–Hoover thermostat and barostat [125,126], with damping coefficients of 100 and 125 fs, respectively. Then, while maintaining a constant temperature for 10 ns in the canonical (NVT) ensemble, the simulation box was deformed until the density of the fluid reached the average density computed from the last 10 ns of the previous NPT simulations. Finally, the system was equilibrated using the final configuration from NVT as the initial configuration for 5 ns in the microcanonical (NVE) ensemble. This way, the desired equilibration state of the system was achieved while avoiding interference with the dynamics of the system [159].

To calculate viscosity, following the equilibration process, further equilibrium molecular dynamics simulations were carried out. Specifically, we used the Green–Kubo approach, which relates the viscosity to the time integral of the stress autocorrelation function [159,160], as in Equation (4.1). Simulations were equilibrated again for 2 ns in the NPT ensemble followed by NVE ensemble production runs. The pressure tensor components were saved every 5 fs for 5 ns at 100 °C and 20 ns at 40 °C during the NVE simulation. An average low-shear viscosity $\langle \eta(t) \rangle$ and its standard deviation $\langle \sigma(t) \rangle$ values were calculated from 20 to 40 NVE trajectories and by applying the Green–Kubo formalism following the time decomposition approach [160] (see Figure B.7). The standard deviation of viscosity was calculated from the multiple NVE trajectories using Equation (4.2). This approach has been used previously in many studies and has been found to provide accurate and reliable viscosity.

$$\eta(t) = \frac{V}{6k_B T}\int_0^\infty \langle P_{\alpha\beta}(t) \cdot P_{\alpha\beta}(0)\rangle dt \tag{4.1}$$

$$\text{Standard deviation, } \langle\sigma(t)\rangle = \left\langle \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\eta(t)_i - \langle\eta(t)\rangle)^2} \right\rangle \tag{4.2}$$

In the Green-Kubo formalism, η is the dynamic viscosity, V is the volume of simulation box, T is the temperature, k_B is the Boltzmann constant, $P_{\alpha\beta}$ is the stress tensor, t is the time, $\langle\eta(t)\rangle$ is the average dynamic viscosity, $\langle\sigma(t)\rangle$ is the standard deviation, and n is the number of NVE trajectories. All six components of stress tensor $P_{\alpha\beta}$ were used to calculate viscosity and used in the denominator of Equation (4.1). Figure 4.2 shows representative results from viscosity simulations of FKV-PAO fluid at 40 °C. In this figure, black and red curves represent the average viscosity and standard deviation, respectively, calculated from 20 NVE trajectories which are shown by 20 blue curves. Each blue curve in Figure 4.2 is made up of 40 blue square symbols which correspond to the viscosity calculated by taking the time integral of stress tensors every 0.5 ns simulation time, as described by Equation (4.1). The trajectories, with their average viscosity and standard deviation curves for all fluids and two temperatures are provided in Figure B.7. An average kinematic viscosity $\langle\nu(t)\rangle$ value from multiple NVE trajectories was calculated by $\langle\eta(t) \times \rho\rangle$. Here, ρ is the density of a fluid calculated from the last 10 ns of NPT ensemble while equilibrating model systems. To simulate accurate viscosity, simulation parameters such as system size, damping and drag coefficients for thermostat and barostat were optimized for the forcefield parameters (see Figure B.5). In addition, the dependence of viscosity on simulation parameters such correlation length, simulation time, and number of NVE trajectory was also studied (see Figure B.6). This vital analysis suggested that the accuracy of viscosity simulations is most highly dependent on correlation length followed by simulation time and that using multiple NVE trajectories is only important to produce reliable viscosity. This approach was used to calculate the Newtonian viscosity of fluids at 40 and 100 °C. These temperatures were chosen to enable calculation of viscosity index as per the ASTM D2270 standard [54] and to reflect the temperature range experienced by lubricants used as hydraulic and automobile fluids.

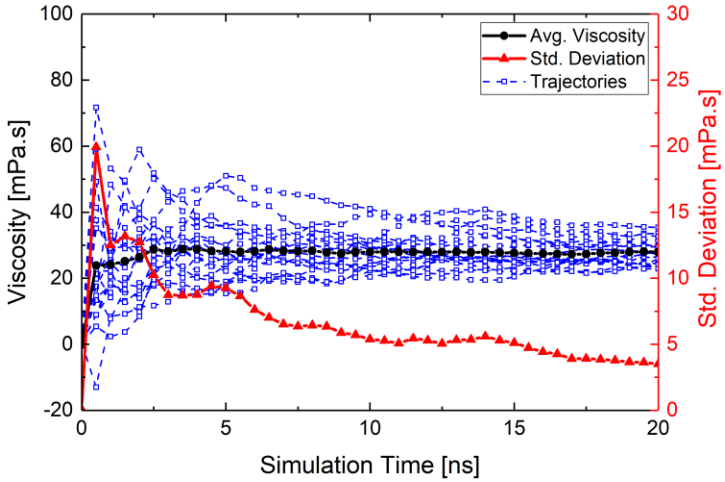


Figure 4.2 Representative viscosity calculation for FKV-PAO fluid at 40°C using the time decomposition approach from 20 NVE trajectories. The blue symbols represent the viscosity obtained from Equation (4.1) by integrating the stress autocorrelation at a given simulation time. The average viscosity and standard deviation are shown by black and red curves, respectively.

To calculate traction coefficient (TC) at different shear rates, loading simulations and then non-equilibration molecular dynamics (NEMD) simulations were carried out. The most common approach to simulating TC is to model the dynamic behavior of fluid molecules confined between two solid surfaces [110,111,113,114,183,184]. In this methodology, solid surfaces are slid with a constant velocity to shear the lubricant, while a constant pressure is applied in the surface-normal direction [110,111,113,114,183,184]. The explicit modelling of the solid surfaces results in a large number of atoms in the simulation which presents a challenge in the case of fluids formulated with large polymer molecules, as in this study. Further, the typical approach requires accurate modeling of solid-liquid interactions, which can be problematic for real engineering materials. Therefore, we developed a simple approach where the pressure and shear are applied directly to the fluid molecules by deforming the simulation box. This approach is appropriate for modeling the physical experiment that was performed in the full film lubrication regime in which traction is due to resistance to shear within the fluid, as opposed to interactions between confining walls. However, a limitation of this approach is that it is not suitable for investigating the effect of interface chemistry.

In the first step of the traction simulations, simulations were carried out to apply a pressure of 0.5, 1.0, or 1.5 GPa. For this, the size of the simulation box was decreased in the z-direction until that the component of stress tensor, or pressure, in the z-direction P_{zz} reached the target value. This approach is equivalent to performing simulations in the NPT ensemble to achieve a target

value of P_{zz}. Then, the simulation with the new size was run in the NVT ensemble at 40 °C for 5 ns followed by the NVE ensemble for 5 ns to relax the system. In the second step, NEMD simulations were carried out at five shear rates, $\dot{\gamma}$, 1×10^7, 1×10^8, 1×10^9, 1×10^{10}, and 1×10^{11} 1/s. In NEMD simulations, the shear rate was imposed on the simulation box by deforming it in the x-direction for 20-300 ns, depending on the shear rate. At the same time, the fluids were thermostatted using a Nosé Hover thermostat [125] and the SLLOD [129,130] equations of motion. Depending on the shear rate, 5–200 ns of time was allowed for the system to achieve steady-state before the normal pressure P_{zz}, shear stress P_{xz}, and velocities of the atoms in the direction of shear u_x were collected for post processing. It was determined that an NEMD simulation reached steady state if the goodness of a linear fit to the average atom velocities was 98% or greater. The average shear rate ($\dot{\gamma} = \partial u_x / \partial z$) was then calculated from the slope of the linear fit to the average atom velocities in the direction of shear u_x and divided by the simulation box length in z-direction L_z.

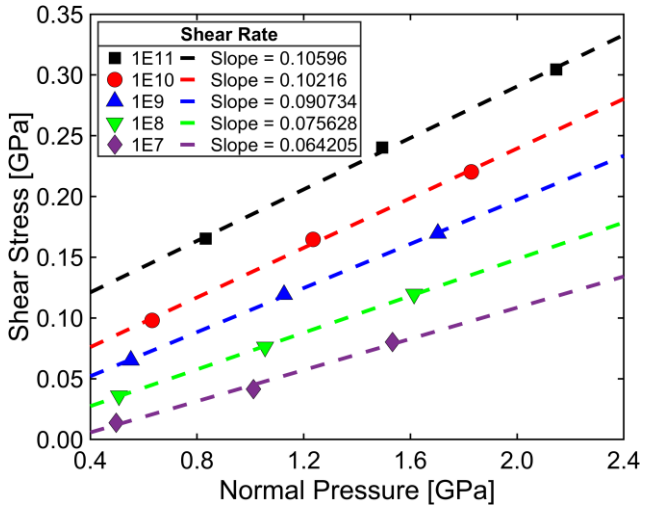
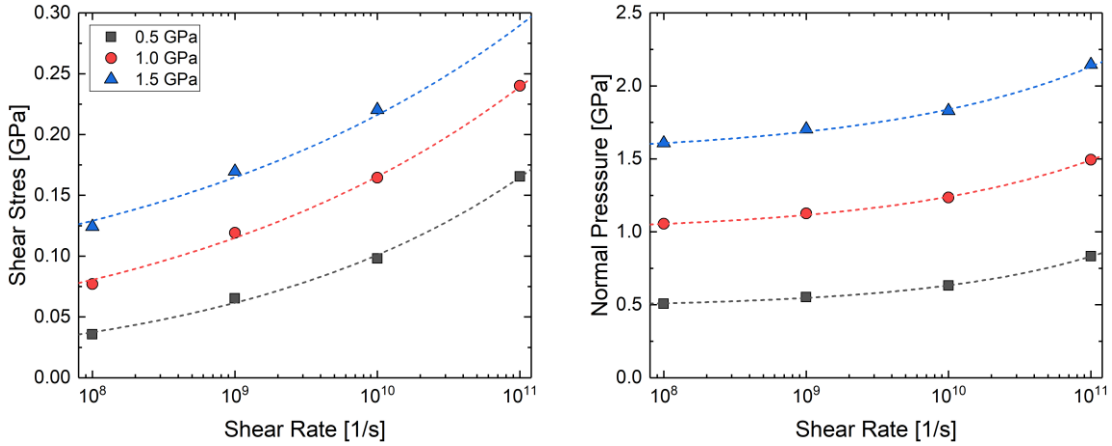


Figure 4.3 Shear stress vs. normal pressure from simulations of FKV-PAO at five different shear rates. The solid black, red, blue, and green symbols represent the simulation data at shear rates of 1×10^7, 1×10^8, 1×10^9, 1×10^{10}, and 1×10^{11} 1/s, respectively, whereas the dashed lines represent the linear fits. The slope of the linear fit at each shear rate is the TC, reported in the legend.

The traction simulations were conducted for each fluid at 40 °C and shear rates of 1×10^7, 1×10^8, 1×10^9, 1×10^{10}, and 1×10^{11} 1/s for each desired normal pressure of 0.5, 1.0, and 1.5 GPa. The average shear stress P_{xz} vs. normal pressure P_{zz} data was fit linearly at each shear rate, and the slope taken as the traction coefficient, as shown in Figure 4.3. The lowest shear rate achievable in the simulations was 1×10^7 1/s. However, the MTM measurements were performed with shear rates of approximately 1.6×10^6 1/s. To extrapolate the

simulation data to this shear rate, we used the relationships between shear stress and shear rate and between normal pressure and shear rate. These relationships are expected to be follow a power-law where the exponent is much larger for normal stress than shear stress [11]. The shear stress and normal pressure were fitted to $P = a + b\dot{\gamma}^c$, where $a, b,$ and c are fitting parameters, as shown in Figure 4.4. The R-squared values of the fits were greater than 99.5% for all the cases. Then, those fitting parameters were used to predict the shear stress and normal pressure at lower shear rates and then the traction coefficients at those shear rates are calculated, as shown in Figure 4.5. This traction simulation approach is very computationally efficient since it does not require using two solid surfaces nor any assumptions about solid-fluid interactions. Linear fits of the shear stress vs. normal pressure for all fluids with their extrapolated traction curves can be found in Figure B.8.



(a) (b)

Figure 4.4 Simulation-calculated (a) shear stress and (b) pressure as a function of shear rate for FKV-PAO at target pressures of 0.5, 1.0, and 1.5 GPa, shown as solid black, red, and blue symbols, respectively. The dashed lines represent power law fits.

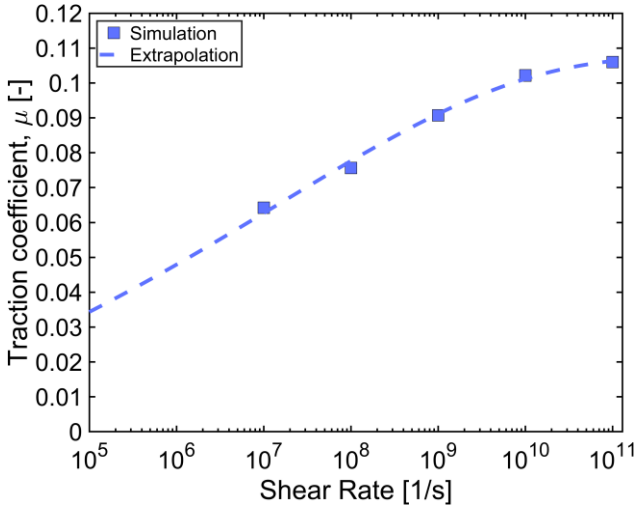


Figure 4.5 Extrapolated traction curve for FKV-PAO generated from the power law fit of shear stress and pressure. The blue dashed line represents the extrapolated curve whereas the solid blue squares represent the traction coefficients directly calculated from the simulation by the slope of linear fit of shear stress to normal pressure at those shear rates.

4.3. Results and Discussion

4.3.1. Validation of Simulation Approach

Figure 4.6 shows the kinematic viscosity measured experimentally and calculated using simulations for all fluids, including the base oil, at 40 and 100°C. The patterned bars are the measured viscosity whereas the solid bars are simulated viscosity at each temperature. The error bars are the standard deviation from multiple experimental measurements or simulation trajectories. Note that the standard deviation for the simulations is higher at 40 °C than at 100 °C because the stress autocorrelation takes greater amount of time to decay at lower temperature due to slower relaxation dynamics, requiring longer simulation durations [185,186]. The effect of simulation duration on viscosity error at different temperatures is illustrated in Figure B.6. Regardless, the simulation results agree with measured values within the reported error.

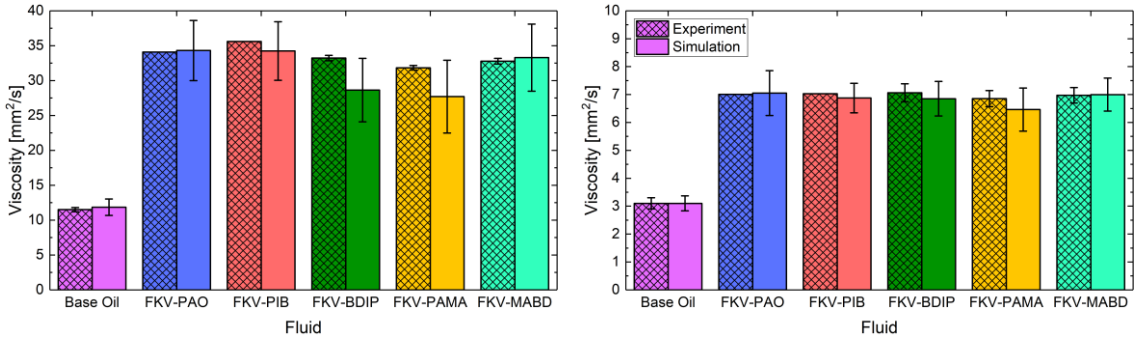
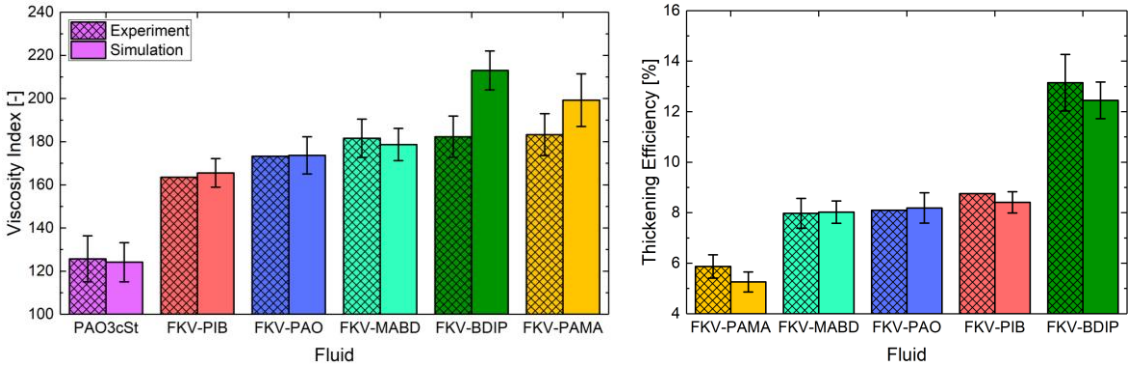


Figure 4.6 Experimentally measured (patterned bars) and simulation calculated (solid bars) kinematic viscosity of the fluids at (a) 40 and (b) 100°C. The error bars represent the standard deviation calculated from multiple experiments and from multiple NVE trajectories (Equation (4.2)) in simulation.

The VI, shown in Figure 4.7(a), was calculated using kinematic viscosity at 40 °C and 100 °C per the ASTM D2270 [54] standard. This test standard states that precision limits cannot be assigned to a calculated VI value. However, we have reported the uncertainty in VI propagated from the standard error of the mean viscosity, as a guide for evaluating the statistical significance of differences between fluids. The order of the mean fluid VIs calculated from experimental data is: Base Oil < FKV-PIB < FKV-PAO < FKV-MABD < FKV-BDIP < FKV-PAMA. The simulation trend of the mean VIs is: Base Oil < FKV-PIB < FKV-PAO < FKV-MABD < FKV-PAMA < FKV-BDIP, the same as in experiments except for the PAMA/BDIP fluids. The simulated VI is higher for FKV-BDIP and FKV-PAMA than measured experimentally. This difference is due to the underprediction of the viscosity of these fluids at 40 °C which can be attributed to the fact that the stress autocorrelation takes longer to decay for larger and more branched molecules due to slower relaxation dynamics, requiring longer simulation durations [185,186]. The average percent difference between experiment and simulation VI for all fluids is 4.98%. However, considering the sensitivity of VI on viscosity of the fluids this is still impressive prediction of VI by MD simulations.



(a) (b)

Figure 4.7 (a) Viscosity index of the fluids calculated using the measured (patterned bars) and simulated (solid bars) kinematic viscosities at 40 and 100 °C. (b) Thickening efficiency of the fluids calculated using the experimentally measured (patterned bars) and simulation-calculated (solid bars) kinematic viscosities at 100 °C. The error bars represent the viscosity error propagated using the VI and TE equations. The error bars represent uncertainty in VI and TE propagated from the standard error of the mean of the viscosities.

The TE at 100 °C was calculated from the measured and simulated kinematic viscosity [55]. Since the fluids had different concentrations of polymer, the thickening efficiencies were normalized by the polymer concentration to enable direct comparison of the fluids, as described by Equation (4.3).

$$\mathrm{TE} = \frac{\nu_{0s} - \nu_{0B}}{\nu_{0B}} \frac{1}{C} \times 100 \tag{4.3}$$

Here, C is the concentration of polymer, ν_{0s} is the Newtonian viscosity of the polymer-containing fluid, and ν_{0B} is the Newtonian viscosity of base oil. The results are shown in Figure 4.7(b). The TE of base oil is not included in these figures since it is the reference and therefore has a TE of zero. In this figure, the patterned bars represent TE calculated using measured kinematic viscosity and the solid bars represent TE calculated using simulated kinematic viscosity. Note that the simulation TE reflects the increase in viscosity due to just the polymer whereas the experimental TE indicates viscosity increase due to both polymer and Anglamol additive package. However, the effect of the additive package on TE is small and the trends are consistent with and without additive, as shown in Figure B.2b. The order of the mean TE in Figure 4.7(b), calculated from either simulation or experimental data, is: FKV-PAMA < FKV-MABD < FKV-PAO < FKV-PIB < FKV-BDIP.

It can be concluded from these comparisons that the properties derived from simulation-calculated viscosity, such as VI and TE, are in good agreement with experimental values. Further, comparing the trends in Figures 4.7(a) and 4.7(b) reveals that a polymer that has higher VI might not also have high TE. Subsequently, we will refer to the mean values of VIs and TEs.

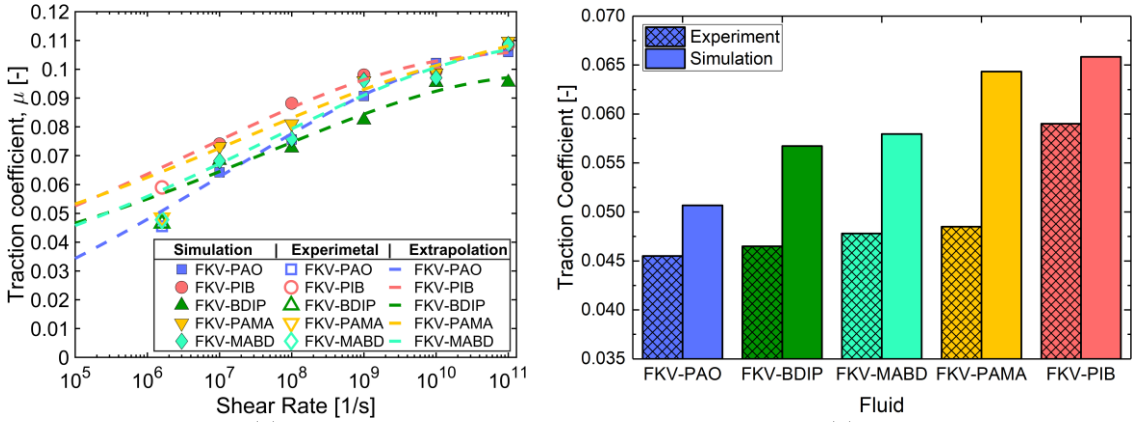


Figure 4.8 (a) Traction coefficient at 40 °C as a function of shear rate. Here, the solid symbols represent the traction coefficients calculated directly from simulations, the hollow symbols represent the measured traction coefficients, and the dashed lines represent the extrapolated traction curve from simulations. (b) Comparison of measured traction coefficient (patterned bars) and the extrapolated simulation traction coefficient (solid bars) at 40 °C and a shear rate of 1.6×10^6 1/s.

Figure 4.8(a) shows the traction coefficient results as a function of shear rate where the hollow symbols are data from experiments and solid symbols are data from simulations. Since the simulations could not directly model the shear rate in the experiments, due to the timescale limitation of the simulation method, we extrapolated the high shear rate traction coefficients from simulations (solid symbols in Figure 4.8a) to the experimental shear rate of 1.6×10^6 1/s. The extrapolation procedure is described in the Materials and Methods section. The dashed lines in Figure 4.8(a) are the extrapolated traction curves. The experimental data and values obtained from extrapolation of simulation results are shown in Figure 4.8(b). The experimental data in Figure 4.8b is the same as the hollow symbols in Figure 4.8a. The overprediction of TC by MD simulations could be due to the fact that TC decreases with increasing film thickness [187] and the simulated film thickness is much smaller than that in experiments. Although the simulations overpredict TC for most fluids, the trend from both experiments and simulations is the same: FKV-PAO < FKV-BDIP < FKV-MABD < FKV-PAMA < FKV-PIB. This good agreement demonstrates that the simulation and extrapolation approaches developed here are useful for modeling the traction coefficient of real lubricants in the full film lubrication regime. Note that the experimental TC includes the effect of the Anglamol additive package while the simulation TC does not. However, as reported in Table B.2 and Figure B.3b, while the additive increases TC, the trends between fluids are the same with and without the additive.

4.3.2. Effect of Structure on Viscosity Index, Thickening Efficiency, and Traction Coefficient

To understand the trends in VI and TE, we related these parameters to the molecular and structural properties of polymers. The two properties to which VI and TE have been correlated are the percent of the molecular weight of the polymer in the backbone (M_{wp}) and the radius of gyration (R_g) of the polymer [51,57,81,82,88,188].

First, it has been proposed that VI is correlated to radius of gyration [51,81,188]. Polymers of higher molecular weight and narrower molecular weight distribution usually have larger VI [67]. A recent experimental study [188] reported that the rate of increase in radius of gyration with temperature was faster for a linear (high M_{wp}) than branched (low M_{wp}) PAMA polymers and that the linear PAMA had higher VI [188]. As shown in Figure 4.9, VI appears to be weakly correlated with M_{wp} and moderately correlated with R_g. However, VI does not vary monotonically for the five fluids studied here with either parameter, indicating that neither parameter can capture the trend on its own. Figure 4.9 shows that, when two polymers have similar M_{wp}, the fluid with the polymer having larger R_g has higher VI. Also, generally, fluids with polymers having larger R_g have higher VI. But, if two polymers have similar R_g, then the fluid with smaller M_{wp} has higher VI. These results suggest that both parameters must be considered to understand the VI trends.

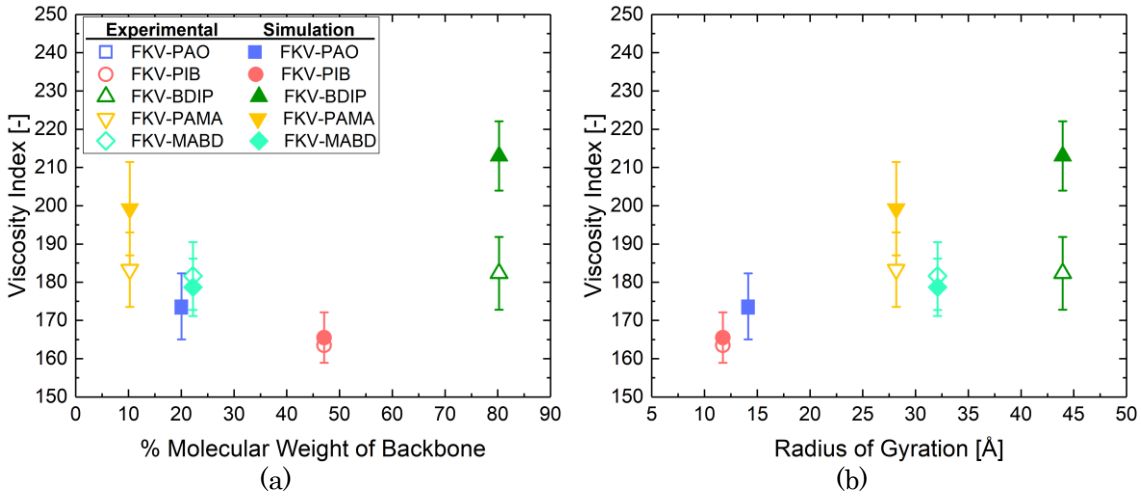


Figure 4.9 Viscosity index of the fluids as a function of (a) percent of the polymer molecular weight in the backbone, and (b) simulation-calculated radius of gyration. The hollow symbols represent viscosity index calculated using measured kinematic viscosity whereas the solid symbols represent viscosity index calculated using simulated kinematic viscosity. The error bars represent uncertainty in VI propagated using the standard error of the mean of the viscosity.

It has been suggested that TE increases with percent of the molecular weight of the polymer in the backbone (M_{wp}) [57,67,68]. One study [57] reported a linear increasing trend of TE with M_{wp} for OCP, PIB, and PAMA polymers. A similar trend of decreasing TE with increased branching (that is lower M_{wp}) in polyethylene polymers was observed in a study of high shear stability VI improvers [67]. For the fluids studied here, Figure 4.10 shows that TE generally increases with M_{wp}, except for the FKV-MABD fluid. This indicates that the linear polymers with higher M_{wp} have a greater thickening effect. However, the increase in TE with M_{wp} is non-monotonic, and TE both increases and decreases with R_g. Between the FKV-PIB, FKV-PAO, and FKV-PAMA fluids, TE decreases as R_g increases, but the opposite trend is observed for FKV-MABD and FKV-BDIP fluid. Thus, the dependence of TE is different than that exhibited by VI: VI generally increases with R_g but varies nonmonotonically correlated with M_{wp}, whereas TE generally increases with M_{wp} but varies nonmonotonically with R_g. Importantly, neither of these two properties cannot individually capture the trends of VI and TE.

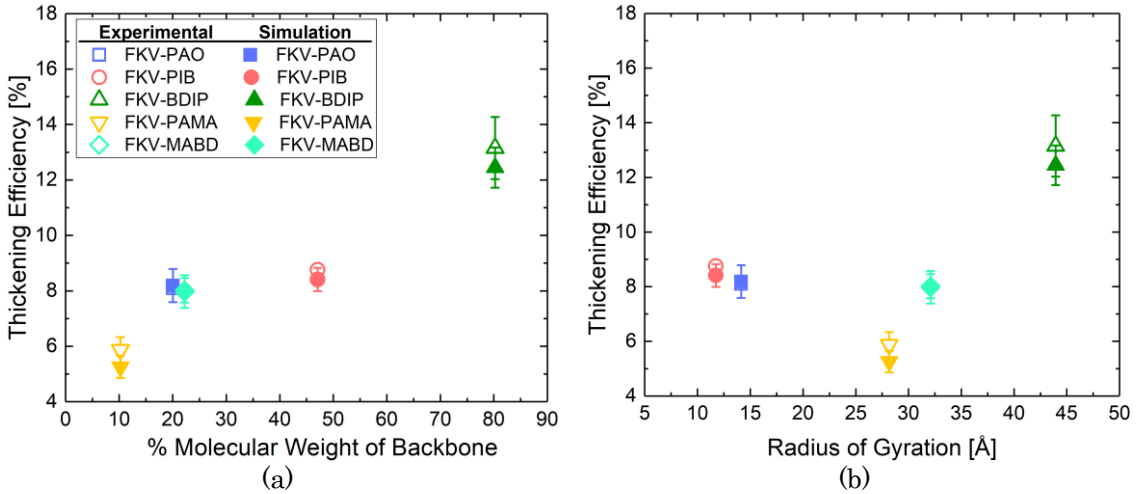


Figure 4.10 Thickening efficiency of the fluids at 100 °C as a function of (a) percent of the polymer molecular weight in the backbone, and (b) simulation-calculated radius of gyration. The hollow symbols represent thickening efficiency calculated using measured kinematic viscosity whereas the solid symbols represent thickening efficiency calculated using simulated kinematic viscosity. The error bars represent uncertainty in TE propagated using the standard error of the mean of the viscosity.

The results in Figure 4.8(b) show that fluids formulated with linear polymers (OCP and BDIP) have lower TC than the fluids formulated with branched polymers (MABD, PAMA, and PAO). A similar trend was previously reported [110] where a higher TC of cyclohexane compared to the benzene was explained by the interlocking of cyclohexane molecules due to their chair conformation as opposed to the plane-shape of the benzene molecules. However, no prior study

investigated the effect of polymer structure on the TC. So, to be consistent with the analyses performed for VI and TE, TC was analyzed as a function of M_{wp} and R_g, as shown in Figure 4.11. Generally, the TC decreases with increasing M_{wp} and R_g of the polymer. However, like VI and TE, the trend is not monotonic. This means that there are likely to be other parameters that correlate to the performance of these lubricants. Therefore, we next evaluated other molecular properties and their potential correlation to VI, TE, and TC.

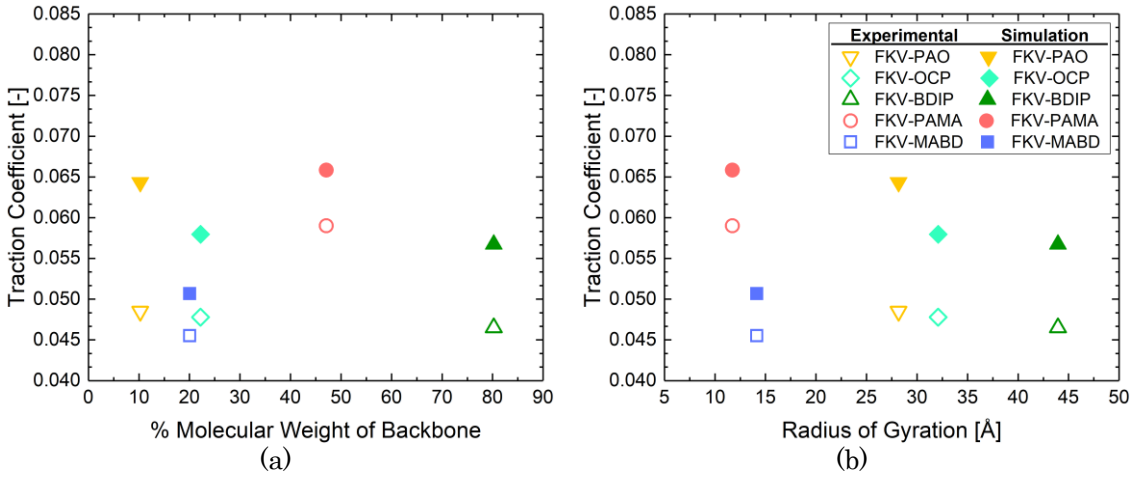


Figure 4.11 Traction coefficient of the fluids at 40 °C and MTM test conditions as a function of (a) percent of the polymer molecular weight in the backbone, and (b) simulation-calculated radius of gyration. The hollow symbols represent TC measured using a MTM machine whereas the solid symbols represent the TC calculated at MTM conditions by extrapolating the traction curve from simulations at high shear rates.

4.3.3. Predictive Models for Viscosity Index, Thickening Efficiency, and Traction Coefficient

To understand the differences between the fluids studied here, we calculated 22 different properties of the polymers. Five of these properties were molecular weights known based on the chemistry of the polymers. These properties are the mass average molar mass (M_w), number average molar mass (M_n), total molecular weight (M_{wt}), percentage molecular weight in backbone (M_{wp}), and molecular weight in backbone (M_{wb}). In addition, we calculated 17 properties based on positions of atoms in the simulations. The simulation-calculated properties are molecules lengths (L_x, L_y, L_z), end-to-end distance (R), radius of gyration (R_g), elements of gyration tensor ($R^2_{g,xx}$, $R^2_{g,yy}$, $R^2_{g,zz}$, $R^2_{g,xy}$, $R^2_{g,xz}$, $R^2_{g,yz}$), eigen values of gyration tensor (λ_x, λ_y, λ_z), and shape parameters (c, b, k). The three computed shape parameters are the asphericity b, the acylindricity c, and the relative shape anisotropy k [189,190]. The subscripts x, y, z refer to the

coordinate axes of the simulation box. More details about each parameter are provided in Section B.10.

We used best subset linear regression to identify models that correlate TE, VI, and TC with individual properties or combinations of the properties. First, the analysis was performed with just the five properties that are functions only of the polymer chemistry (M_w, M_n, M_{wt}, M_{wp}, and M_{wb}). Then, the analysis was repeated with all 22 properties, including both properties that are functions of only chemistry and those calculated in the simulations (L_x, L_y, L_z, R, R_g, $R^2_{g,xx}$, $R^2_{g,yy}$, $R^2_{g,zz}$, $R^2_{g,xy}$, $R^2_{g,xz}$, $R^2_{g,yz}$, λ_x, λ_y, λ_z, c, b, and k). The best models were selected out of more than 2 million possible combinations of these properties based on the R-squared value, standard error, and variance inflation factor (VIF) of the fit. Note that an ideal model would have an R-squared value of 1, a VIF of 1.0 and std. error of 0.0. Some models included only properties that are known based on the chemistry of the polymer whereas others included the best combination of properties out of all 22 properties. All structural parameters used in development of the VI and TE models are reported in Table B.10.

The two best models for VI are given in Equations (4.4) and (4.5). Model A is a 2-term model based only on the first five properties, and Model B is a 3-term model based on all 22 properties. Model A is a function of molecular weight of backbone and percentage of molecular weight in backbone ($M_{wp} = M_{wb}/M_{wt} \times 100$) of the polymer. Model B is a function of number avg. molar mass, molecule length, and relative anisotropy ratio of the polymer molecules. Since VI is calculated using the kinematic viscosities at 40 and 100 °C, the polymer properties were averaged over the two temperatures. The fit parameters have base units of kilogram as the unit of mass and angstrom as the unit of distance. The coefficients and values of the model selection criteria are given in Table 4.2. The coefficient of determination (R-squared) and VIF were higher, and the standard error (std. error) lower with Model B, indicating the model with the simulation parameter is a better predictor of VI. The magnitude of the coefficients indicates the degree to which each parameter is correlated to VI and the sign of the coefficient indicates and whether the correlation is positive or negative. For instance, in Model A, the coefficient for M_{wb} is 2.630 and for M_{wp} is -0.398. This means, M_{wb} is positively correlated whereas M_{wp} is negatively correlated (that is, if M_{wb} increases, VI will increase and if M_{wp} increases, VI will decrease) but M_{wb} is approximately 7 times more correlated to VI than M_{wp}. The coefficients of Model B can be interpreted similarly.

$$\text{A: VI} = a_0 + a_1 M_{wb} + a_2 M_{wp} \tag{4.4}$$

$$\text{B: VI} = a_0 + a_1 M_n + a_2 L_y + a_3 k \tag{4.5}$$

The two best models for TE at 100 °C are given in Equations (4.6) and (4.7). Model A, which was developed from only the first set of parameters, described TE as a function of total molecular weight of polymer and molecular weight of backbone of the polymer. When both sets of parameters were used, Model B was obtained which includes total molecular weight, end-to-end distance, and acylindricity of the polymer molecules. The polymer properties in Model B are from only the viscosity simulations at 100°C since the model was developed to predict TE at 100 °C. The coefficients and values of the selection criteria of these models are given in Table 4.2. The coefficients of TE models indicate that TE is negatively correlated to total molecular weight of the polymer. However, TE is positively correlated to molecular weight of backbone of the polymer like VI. This commonality in models of VI and TE suggest that a polymer could be designed to provide fluid with improved VI as well as improved TE (that, is an amount of polymer required to achieve a desired viscosity), as it the case, for the BDIP polymer which provided the highest TE and the second highest VI amongst all polymers. From the developed VI and TE models, it can be said that both are highly dependent on the molecular weight because the viscosity of lubricants is highly dependent of the molecular weight [191].

$$\text{A: \%TE} = a_0 + a_1 M_{wt} + a_2 M_{wb} \tag{4.6}$$

$$\text{B: \%TE} = a_0 + a_1 M_{wt} + a_2 R + a_3 c \tag{4.7}$$

The three best models for TC are given in Equations (4.8), (4.9), and (4.10). There was no model based on only the first set of parameters which could give an R-squared value above 5%, so a two-term based model on both sets of parameters was developed and presented as Model A. Model A described TC as a function of molecular weight of the backbone and relative anisotropy ratio of polymer. When both sets of parameters were used to develop a three-term model, Model B was obtained, which includes mass average molar mass, percentage of molecular weight in backbone, and relative anisotropy ratio of the polymer. Neither of these models was sufficiently accurate, so Model C was identified which is a five-term model that has terms from both the sets of parameters. The terms in Model C include total molecular weight, percentage molecular weight in backbone, molecules length in the z-direction L_z, square of radius of gyration in zz-direction $R^2_{g,zz}$, and again the relative anisotropy ratio k of the polymers. These TC models were developed using both experimental and simulated TC. For the simulation data, the polymer properties at each shear rate were averaged over the three pressures simulated. The coefficients and values of the selection criteria of these models are given in Table 4.2. Note that the R-squared value is considerably lower for the TC than for the VI and TE models due to the fact that TC data from a wide range of shear rates (1.6×10^6 - 1×10^{11} 1/s) was used in the fitting. As shown in Figure 4.10(a), TC

increases non-linearly with shear rate whereas all the empirical models to which the data is fit are linear.

The coefficients for Model C in Table 4.2 show that the percentage of molecular weight in the backbone of the polymer M_{wp} and polymer length in the direction of normal pressure L_z are negatively correlated whereas total molecular weight of polymer M_{wt}, square of radius of gyration in zz-direction $R^2_{g,zz}$ and relative anisotropy ratio k are positively correlated to TC. Of the five terms in Model C, the relative anisotropy ratio k is the most correlated and the square of the radius of gyration in the zz-direction $R^2_{g,zz}$ is the least correlated to TC. The signs of these coefficients also make physical sense. For instance, the TC of each fluid increases with shear rate whereas L_z decreases and k increases with shear rate (see Table B.11), so the sign of coefficient of L_z is negative and of k is positive. All structural parameters used in the development of the TC models are listed in Table B.11.

$$\text{A: TC} = a_0 + a_1 M_{wb} + a_2 k \tag{4.8}$$

$$\text{B: TC} = a_0 + a_1 M_w + a_2 M_{wp} + a_3 k \tag{4.9}$$

$$\text{C: TC} = a_0 + a_1 M_{wt} + a_2 M_{wp} + a_3 L_z + a_4 R^2_{g,zz} + a_5 k \tag{4.10}$$

Table 4.2 Coefficients and selection criteria for predictive models of VI, TE, and TC. Kilogram as the unit of mass and angstrom as the unit of distance were considered for these parameters.

Model		A	B	C
VI	Coefficients	$a_0 = 180.143$ $a_1 = 2.630$ $a_2 = -0.398$	$a_0 = 158.841$ $a_1 = 0.341$ $a_2 = 0.573$ $a_3 = -29.264$	No Need
	Selection Criteria	$R^2 = 0.932$ $VIF = 2.688$ $std.\ Error = 3.123$	$R^2 = 1.000$ $VIF = 7.119$ $std.\ Error = 0.004$	No Need
TE	Coefficients	$a_0 = 8.300$ $a_1 = -0.131$ $a_2 = 0.534$	$a_0 = 11.099$ $a_1 = -0.078$ $a_2 = -0.108$ $a_3 = 0.008$	No Need

	Selection Criteria	$R^2 = 0.997$ $VIF = 1.106$ $std.\ Error = 0.204$	$R^2 = 1.000$ $VIF = 2.510$ $std.\ Error = 0.018$	No Need
TC	Coefficients	$a_0 = 2.46 \times 10^{-2}$ $a_1 = -42.87 \times 10^{-4}$ $a_2 = 11.40 \times 10^{-2}$	$a_0 = 5.68 \times 10^{-2}$ $a_1 = -36.14 \times 10^{-4}$ $a_2 = -10.73 \times 10^{-4}$ $a_3 = 17.35 \times 10^{-2}$	$a_0 = 10.70 \times 10^{-2}$ $a_1 = 8.02 \times 10^{-4}$ $a_2 = -5.43 \times 10^{-4}$ $a_3 = -23.47 \times 10^{-4}$ $a_4 = 1.80 \times 10^{-4}$ $a_5 = 6.29 \times 10^{-4}$
	Selection Criteria	$R^2 = 0.629$ $VIF = 2.181$ $std.\ Error = 0.013$	$R^2 = 0.746$ $VIF = 3.396$ $std.\ Error = 0.011$	$R^2 = 0.824$ $VIF = 18.756$ $std.\ Error = 0.009$

4.3.4. Predictive Model Validation

To validate the predictive models, a new fluid was physically formulated. The new fluid, called F-ROMP, comprised a 5.8 wt.% new polymer ROMP ($C_{704}H_{1410}$), shown in Figure 4.12, with 48.1 wt.% Nexbase 3030 base oil and 39.6 wt.% Nexbase 3043 base oil. The Nexbase 3030 and 3043 base oils are 3.0 and 4.3 mm^2/s oils, respectively. The physically blended F-ROMP also had 6.5 wt.% Anglamol additive, like the original five fluids. The polymer in F-ROMP is neither linear like PIB and BDIP nor branched like PAO, MABD, and PAMD. The degree of branching in ROMP polymer is in between of these two sets of polymers. The number average molecular weight and polydispersity index of ROMP polymer were 10 kg/mol and 1.5, respectively. The measured kinematic viscosity of F-ROMP was 5.6 mm^2/s and 25.5 mm^2/s at 100 and 40 °C, respectively. VI, TE, and TC were measured using experimental data for this fluid (values in the Supporting Information).

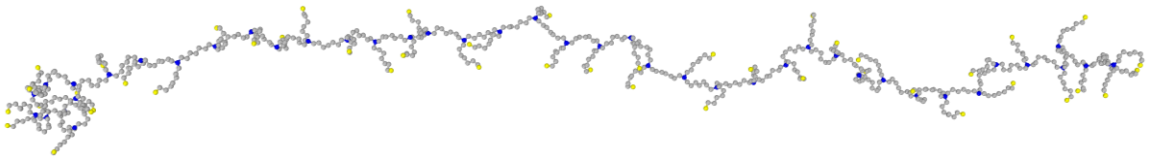


Figure 4.12 Structure of ROMP polymer in the F-ROMP fluid. The sphere colors represent the following pseudoatoms: yellow-CH_3, silver-CH_2, and blue-CH.

Next, the new fluid was created in the simulations. Since we did not know the chemistry of the Nexbase oils, we replaced Nexbase 3.0 cSt with the previously formulated PAO 3.0 cSt (see Table B.3) and Nexbase 4.3 cSt with a newly identified blend of 84.3 wt.% of PAO 4 cSt and 15.7 wt.% of PAO 8 cSt (the viscosity of this blend of PAO 4 cSt and PAO 8 cSt is 4.28 cSt at 100°C = PAO4.3). The resultant PAO base oil has the same viscosity as the blend of

Nexbase base oils. Therefore, the F-ROMP fluid model was virtually modeled in Material Studio software by mixing 5.8 wt.% of F-ROMP polymer, 51.7 wt.% of PAO3, and 42.5 wt.% of PAO4.3. Simulations were run with the new fluid to calculate the 17 structural parameters used in the simple models at 40 and 100 °C.

The VI of F-ROMP was calculated using Equation (4.5), TE was calculated using Equation (4.7), and TC was calculated using Equation (4.10) with the fit parameters in Table 4.2. The model predicted VI and TE are compared to experimentally measured data in Figures 13(a) and 13(b). As expected, the model predictions are very accurate, particularly the trends and magnitudes, for the five fluids used to fit the model parameters. For the F-ROMP, the model predicted TE very accurately and only slightly overpredicts VI. Importantly, the model accurately predicts the relative position of F-ROMP in the increasing VI trend. For TE, the model predicted both the trend and absolute value of TE for all the fluids well, including the test fluid F-ROMP. This degree of matching is impressive since the models were developed for a different base oil than used in the F-ROMP formulation.

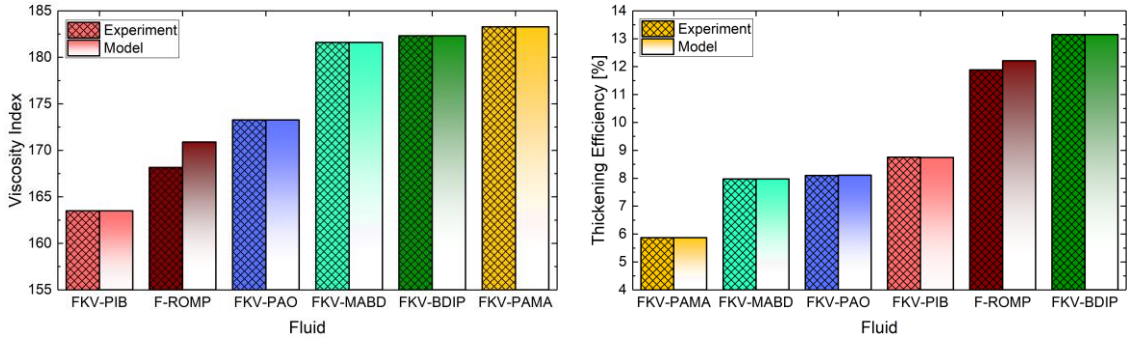
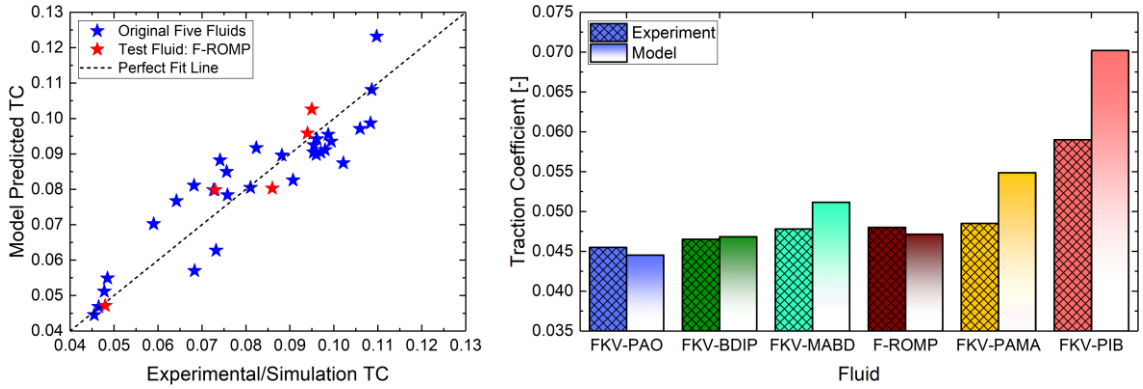


Figure 4.13 (a) VI and (b) TE at 100 °C for all fluids, including new test fluid F-ROMP, from experiments (patterned bars) and the empirical models fit to just the first five fluids (bars with white-to-solid gradient).

The simulated TC data and extrapolated traction curve for F-ROMP are reported in the Appendix Section B.12. The TCs from simulation, experiment, and predicted by Model C at a range of shear rates are shown in Figure 4.14(a). In this figure, the blue dashed line is a perfect fit, and the blue and red asterisks are the TC for the five original fluids and F-ROMP, respectively, both from the experiments and simulations. To compare the model-predicted TC to experimental values, the original five fluids were simulated at shear rate consistent to MTM test conditions, that is 1.6×10^6 1/s, whereas the F-ROMP was simulated at shear rate of 1.9×10^6 1/s. Then, using the coordinates of the atoms, the polymer properties in Model C were calculated to predict TC at the experimental shear rate. The results are shown in Figure 4.14(b). While the

model accurately predicts the trend of the original five fluids and magnitude of the test fluid F-ROMP, it does not correctly place the test fluid F-ROMP in the increasing TC trend. This error may be due to inaccuracies in estimation of the MTM shear rate, the non-linearity of the TC with shear rate, the difference between the experimental and simulated film thickness, as well as, for F-ROMP, the difference in the simulated and actual base oil. However, the accuracy of the TC model is still impressive and, notably, the model can predict TC over wide range of shear rates since it was developed from experimental TCs at 1.6×10^6 and simulated TCs at shear rates between 1×10^7 and 1×10^{11} 1/s.



(a) (b)

Figure 4.14 (a) Model predicted traction coefficients for all fluids at high shear rates to directly compare to simulation calculated. The blue dashed line is a perfect fit line, blue asterisks represent simulation-calculated TC for the five fluids used to develop the model and red asterisks represent the TC predicted using model and the simulation data for the test fluid F-ROMP. (b) TE at 40 °C for all fluids, including new test fluid F-ROMP, from experiments (patterned bars) and the empirical models fit to just the first five fluids (bars with white-to-solid gradient).

4.4. Conclusion

In this study, MD simulations, experimental measurements, and empirical modeling were used to study the effect of chemistry and structure of polymers on three lubricant performance metrics. Five fluids of similar viscosity were formulated with PAO, PIB, BDIP, PAMA, and MABD polymers in PAO base oil of 3 mm^2/s. These polymers cover a broad range of molecular weight, chemistry, and degree of branching. Both kinematic and dynamic viscosity were measured using a viscosimeter and a rheometer at 40 and 100 °C, and the VI and TE calculated. The TC was measured using an MTM. The measured data was then used to validate the MD simulations. For the viscosity simulations, various simulation parameters were optimized and the dependency of viscosity simulation approach on correlation length, simulation time, number of trajectories were investigated. Both the kinematic and

dynamic viscosity of fluid at both temperatures were in very good agreement with the experimental data. Then, an accurate and computationally efficient approach was presented to simulate TC. In this approach, the confined molecules were sheared by directly deforming the simulation box instead of modeling solid surfaces. The computational efficiency and accuracy of the method were demonstrated here, although further assessment is required to assess the influence of system size and aspect ratio. A simple approach was also presented to extrapolate simulated traction data to experimentally accessible shear rates. The resultant TCs were slightly higher than experimental values, but the trends were consistent.

Comparing the different fluid formulations, the following trends were observed:

- VI: FKV-PIB < FKV-PAO < FKV-MABD < FKV-BDIP < FKV-PAMA,
- TE: FKV-PAMA < FKV-MABD < FKV-PAO < FKV-PIB < FKV-BDIP,
- TC: FKV-PAO < FKV-BDIP < FKV-MABD < FKV-PAMA < FKV-PIB.

From this, was concluded that the branched polymers provided higher VI, while linear and longer polymer provided higher TE. However, the dependence of TC on branching was not clear. For example, PIB is both more linear and smaller than the other polymers but still had the highest TC, whereas PAO is both small and branched but provided the lowest TC amongst all the polymers. The difference in TC was not clear, it could also be since the difference in chemistries of PAO, BDIP, PAMA, and MABD is comparatively small. The difference in VI, TE, and TC of the fluids were studied in terms of the two most used structural parameters, percentage molecular weight of the polymer in backbone and radius of gyration of the polymer. This investigation showed that VI is somewhat correlated to radius of gyration and TE is highly correlated to the percentage of molecular weight in the backbone of the polymer. However, TC was very weakly correlated with either of structural parameters. Further, although some correlations were observed, neither parameter could capture all trends exhibited by the five fluids studied here. Therefore, to capture the effects of multiple molecular parameters, we used the simplest machine learning tool possible, i.e., multiple linear regression.

To quantify the molecular origins of the VI, TE, and TC, 22 different structural parameters were calculated, including five properties from various molecular weights based on chemistry and 17 properties from the positions of atoms in the MD simulations. Empirical models were developed based on these parameters. It was found that better models, i.e., more accurate predictions, could be achieved using the simulation properties. The best models were

evaluated by comparing their predictions to experimental data for a fluid not included in the model development. The models accurately predicted both the magnitude and trend of VI and TE for all the fluids, including the test fluid that was not used in model development. The TC model accurately predicted the trend of all fluids except the test fluid and accurately predicted the magnitude of the test fluid TC. The limitations of the model are primarily attributed to the fact that only five fluids were included in the parameterization and that the test fluid had a different base oil than that of the fluids used in model development. The next step could be to expand the data set and further develop the mathematical models to correlate dynamic descriptors from the simulations to lubricant performance metrics.

Overall, we have demonstrated use of molecular simulations to accurately predict VI, TE, and TC of lubricant formulations with commercial relevance. We further used the simulations to develop empirical models based on properties that could be calculated using relatively short simulations. Adding more data to the models, e.g., different polymer concentration, different base oil, etc. would make the model more accurate. Ultimately, we believe this can be an approach to accurate prediction of structure-property-function relationship that can guide design of new lubricants or additives.

Chapter 5. PyL3dMD: Python LAMMPS 3D Molecular Descriptors Package

5.1. Introduction

Molecular dynamics (MD) simulations are used to study the physical and chemical properties of materials [192]. There are many software and packages for performing MD simulations, including LAMMPS [193], AMBER [194], GROMACS [195], CHARMM [196], DESMOND [197], Materials Studio [198], NAMD [199], and QuantumATK [200]. LAMMPS is one of the most widely used open-source packages for MD simulations, attracting particular interest in the scientific research community due to its stability, flexibility, functionality, and responsive community support [192,193,201].

Although scientific studies and the development of novel materials have benefited from MD simulations using LAMMPS, the efficiency and computational cost of these methods still limits the size and time scale of the materials and processes that can be studied. In recent years, with the development of artificial intelligence, significant interest has arisen in machine learning (ML) as a design tool. This approach to predicting material properties is called quantitative-structure-property-relationship (QSPR) modeling and is becoming an essential technology in a wide variety of research fields because of its computational efficiency, scalability, robustness, and predictive ability.

QSPR modeling is building mathematical relationships between material properties and molecular descriptors of the molecules that compose that material. Molecular descriptors are quantitative representations of physical, chemical, or topological characteristics of molecules that summarize our knowledge and understanding of molecular structure and activity from different aspects [202,203]. Molecular descriptors play a fundamental role in chemistry, pharmaceutical sciences, environmental protection policy, health research, and quality control. QSPR models based on molecular descriptors have been widely used in pharmaceutical [204,205] industries and predicting the biological [206] and physiochemical [207–209] properties of molecules.

There are currently thousands of molecular descriptors, which can be classified into three broad categories: 1D, 2D, and 3D descriptors, where D stands for dimension(s). Simple molecular descriptors derived by counting atom types or structural fragments in the molecule are called constitution or 1D descriptors. Descriptors derived from algorithms applied to a topological representation (molecular graph) are called topological or 2D descriptors. Lastly, there are

molecular descriptors derived from geometrical representations of molecules called geometric or 3D descriptors [210]. A descriptor can be simple, like molecular volume, which encodes only one feature of a molecule, or can be complex, like GETAWAY [211], which encodes multiple features - geometry, topology, and atom-weights assembly of a molecule.

Various open-source and proprietary software packages have been developed to calculate descriptors, including PaDEL [212], BlueDesc [213], Dragon [210], RDKit [214], CDK [215], Cinfony [216], Chemopy [217], ChemDes [203], BioJava [218], BioTriangle [219], Bioclipse [220], Propy [221], PyDPI [222], RepDNA [223], CDK-Taverna [224], Protr/ProtrWeb [225], JCompoundMapper [226], ChemmineR [227], and Rcpi [228]. In these packages, a molecular structure must be provided to calculate descriptors for a given molecule. The most common format of input for descriptor calculations is Simplified Molecular-Input Line-Entry System (SMILES) [229].

Although SMILES is easy and fast for calculating simple 1D and 2D descriptors by simple operations such as counting atom types or chemical fragments, it does not contain the information necessary for calculating 3D descriptors, such as the time-dependent geometries of molecules. For calculating 3D descriptors, the available molecular file formats include Sybyl MOL2 files (.mol, .ml2, mol2) by Tripos, Inc., Sybyl Molfiles (.sm2) by ChemOffice, CambridgeSoft Corp., Multiple SD files (.sdf) by Molecular Design Ltd., HyperChem files (.hin) by Hypercube, Inc., MacroModel files by Schrodinger [210]. These files contain geometric information for one molecule and one time instance, so the descriptors are calculated for only a single molecule in a given configuration.

In contrast, MD simulations provide geometric information for multiple molecules and time frames. However, this information cannot be readily used by the currently available packages for calculating 3D descriptors. Specifically, with existing packages, the geometric information for each molecule at each timestep must be converted into the file format required by each package, resulting in the generation of a large number of unnecessary files and computational inefficiency. In addition, most existing packages heavily focus on the calculations of simple 1D and 2D descriptors. Therefore, there is a need for a tool specifically oriented to 3D descriptors and MD simulations.

Another challenge is that available descriptor calculation packages do not directly accept the file and data structures that are output from typical MD simulations, for example, input data files (.lmp) and output trajectory files (.lammpstrj) from LAMMPS. Therefore, this paper presents a new Python package, PyL3dMD, where Py stands for Python, L for LAMMPS, 3d for 3-dimensional, and MD for molecular dynamics/descriptors. PyL3dMD can

calculate 1982 3D descriptors directly using the LAMMPS input data and output trajectory files from a simulation.

5.2. General Features

PyL3dMD is written in Python, which is readily available and allows for readable code. PyL3dMD uses *argparse* to provide a command-line interface to pass inputs to the package, which allows multiple optional input parameters to be passed efficiently and for automation using shell scripts.

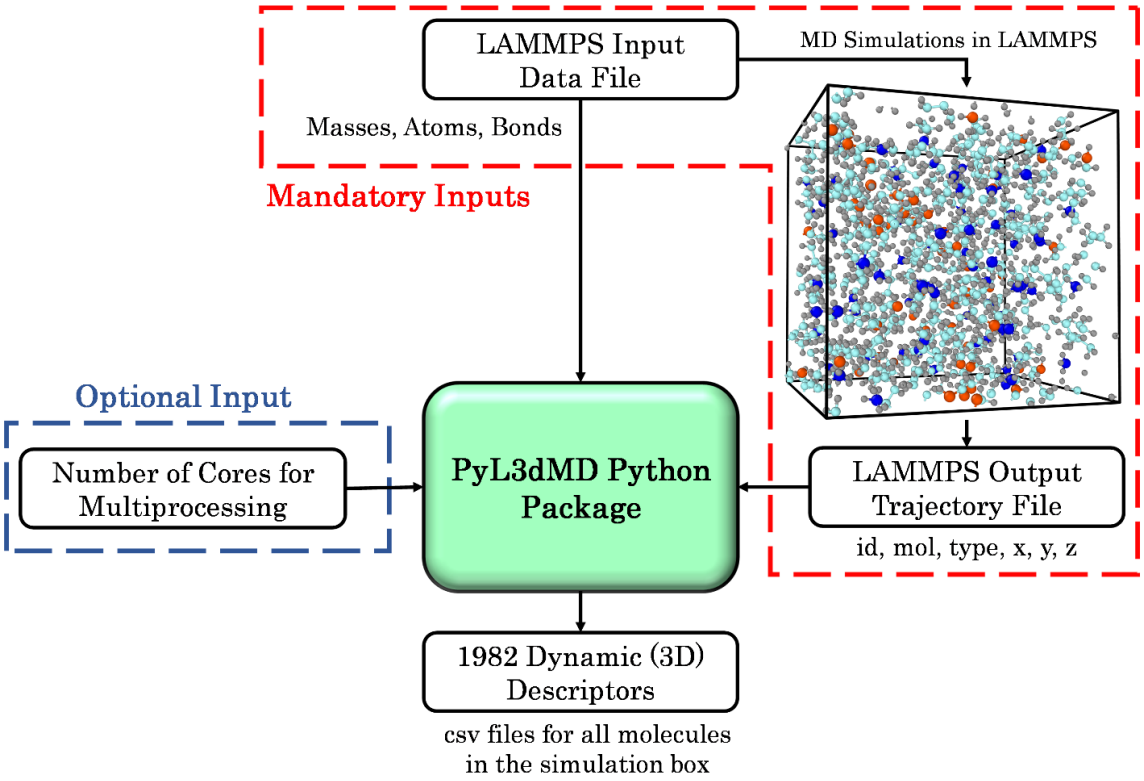


Figure 5.1 Overview of the package and its usage.

PyL3dMD is coded in a module-oriented manner, where each type of descriptor is represented by its module. Furthermore, each module contains a driver function/module used to calculate the respective set of descriptors. This allows PyL3dMD not only to be used through the command line but also to be easily integrated into scripts for user-oriented analyses. Example scripts are provided in the supplementary materials. Owing to the modular structure of PyL3dMD, extensions or new descriptors can be implemented quickly and easily without time-consuming and complex modifications to the source code. To add a new descriptor, users only need to create a new function for that

descriptor and call it in the main module. PyL3dMD has the following dependent Python packages: Math, Pandas, NumPy, Multiprocessing, and Time, which must be installed before using PyL3dMD. All these packages are preinstalled and therefore PyL3dMD does not require any third-party package to be installed.

PyL3dMD is an intelligent parsing of the LAMMPS input data and output trajectory files. The order of sections (atoms, bonds, angles, dihedrals, etc.) in the LAMMPS input data file can vary. PyL3dMD automatically detects each section to parse relevant information. In addition, LAMMPS allows users to export/write many parameters in any order (that is, in any column) to the output trajectory file. PyL3dMD automatically determines the location of the parameters relevant to the calculation and sorts the coordinates using atom and molecule IDs, which allows users flexibility in the format of the output from LAMMPS.

Another feature of PyL3dMD is that it works with any simulation box size with periodic boundaries, and for any box shape for which all sides of the box are perpendicular to one another, e.g., cubic and orthogonal boxes. LAMMPS allows users to export wrapped and unwrapped x, y, and z atom coordinates. Although most 3D descriptors are calculated using unwrapped coordinates of atoms, if needed, PyL3dMD automatically converts wrapped coordinates to unwrapped coordinates before any calculations.

Since the PyL3dMD package requires bond, angle, and dihedral information to calculate 3D topological and connectivity descriptors, the package only works for force fields that have this information in the LAMMPS input data file. The package has been tested for non-reactive force field – OPLS [122] and TraPPE-UA [153], but should work for most non-reactive force fields that have Masses, Atoms, Bonds sections in the LAMMPS input data file and have id (ID of atoms), mol (ID of molecules), type (type of atoms), and 3D coordinates (x, y, z) in the LAMMPS output trajectory file. PyL3dMD in its current form does not work for reactive force fields such as ReaxFF [230]. PyL3dMD can be used with either united-atom (UA) or all-atom (AA) representations. The package adds and removes hydrogen atoms depending on the descriptor type. In addition, it can be used for systems with single or multiple molecule types. For instance, PyL3dMD can calculate descriptors for a fluid that was formulated by blending different molecules, such as polymer and solvent molecules [231].

An important consideration is the efficiency of the calculation, especially when the system consists of many molecules and trajectories containing data from many time frames. For computational efficiency, PyL3dMD uses matrix algebra wherever possible, instead of for and while loops. PyL3dMD can

perform parallel computation for faster estimation of descriptors for all time steps and molecules. If run on desktop computers, PyL3dMD can automatically detect the number of cores available in the system and use them for multiprocessing.

To calculate descriptors, the PyL3dMD package takes three inputs, two of which are mandatory, and one is optional, as shown in Figure 5.1. The two mandatory inputs are the LAMMPS input data and output trajectory files, with their file locations. The LAMMPS input data file must have masses, atoms, bonds, angles, and dihedrals sections, including charges of the atoms. The LAMMPS trajectory file must have atom id, molecule id, atom type, x (or xu), y (or yu), and z (or zu). Here, x, y, and z are coordinates of wrapped atoms, whereas xu, yu, and zu are coordinates of unwrapped atoms. The optional input is the integer number of cores (default is maximum available processor) for multiprocessing. It is recommended that users provide this optional input argument for faster calculation of molecular descriptors.

On successful execution, PyL3dMD generates a comma-separated values (csv) file for each molecule in the simulation box consisting of molecular descriptors for all time frames. The naming convention for these csv files is `molecule_molID.csv`, where the `molID` is the ID of a molecule obtained from the LAMMPS output trajectory file. This way users have freedom to further post-process the calculated descriptors as needed, for example, averaging descriptors over time or multiple molecules.

The PyL3dMD code has several limitations currently. First, PyL3dMD assumes that the LAMMPS simulation is using the real unit system (mass = grams/mole, length = Angstroms, time = femtoseconds, energy = kcal/mole, temperature = Kelvin, pressure = atmospheres, electric field = volts/Angstrom, density = gram/cm^3). However, this may not be an issue for QSPR modeling that is based on the relative values of descriptors for different molecules. The only system property that is calculated is density, so the units of density may need to be converted. Second, PyL3dMD assumes that all the sides of the simulation box are perpendicular to one another when unwrapping the coordinates if the user provides wrapped coordinates. Third, it assumes that the atom style is full when importing the atoms section from the LAMMPS input data file. Finally, since this package calculates descriptors for each molecule separately over multiple time steps, it might require a large amount of memory to store arrays, lists, and dictionaries. In this case, the package should be run on a high-performance computing (HPC) cluster.

5.3. Molecular Descriptors

PyL3dMD can compute 1982 3D descriptors in the current version, including structural and physicochemical descriptors, as summarized in Table 5.1. The complete descriptor list is given in the manual provided in supplementary materials. The governing equations for the descriptors are also provided in the manual.

Table 5.1 Summary of the molecular descriptors calculated using the PyL3dMD package, with the number of descriptors for each descriptor group and type.

Descriptor Group	Descriptor Type	Number of descriptor(s)
3D Topological	• Topology descriptors	18
3D Connectivity	• Connectivity indices	9
Geometric	• Inertia index • Gyration index • Gravitation index • Molecular volume • Shadow indices • Plan of best fit score • Miscellaneous	16 15 6 2 3 2 43
3D RDF	• RDF descriptors	240
3D Autocorrelation	• Moreau-Broto autocorrelation • Moran autocorrelation • Geary autocorrelation	210 210 210
Combination	• 3D-MoRSE • GETAWAY • WHIM • CPSA	240 611 112 30
Properties	• Density (system property) • Dipole moment (molecular property)	1 4

Figure 5.2 shows an example of a Python script that uses the PyL3dMD package for calculating descriptors for a sample model with parallel computation with 4 CPU cores.

```
from PyL3dMD import PyL3dMD

if __name__ == "__main__":
    datafilename =  'Sample.lmp' # your data file name
    dumpfilename = 'Sample.lammpstrj' # your data file name
    SetNumberOfCore = 4 # Number of CPUs for parallel computation
    program = PyL3dMD.PyL3dMD(datafilename, dumpfilename, numberofcores=SetNumberOfCore)
    program.start()
```

Figure 5.2 Screenshot of a sample Python script of using PyL3dMD.

The time required for the code to run is a function of the computational resource used (number of cores), system size (number of atoms in the simulation box), and the duration of the MD simulation (number of time frames). The code was tested for a simulation of 10 molecules with 20 atoms per molecule with 1001 time frames in the LAMMPS output trajectory file. The machine used for this analysis had the following configuration: Intel i7-10700 CPU with 8 cores (16 processors), 32 GB RAM, and Windows 10 desktop. When multiprocessing with 8 cores was used, PyL3dMD took on average 58 minutes to calculate all descriptors whereas the single core took 456 minutes.

5.4. Conclusions

Currently, there are several commercial and open-source software tools for calculating molecular descriptors, but none of them are compatible with MD simulation tools such as LAMMPS. Here, an open-source Python-based 3D molecular descriptors calculation tool, PyL3dMD, was developed, which is compatible with the formats of LAMMPS input data and output trajectory files. PyL3dMD has been published on GitHub, PyPi, and Conda under the GNU General Public License. PyL3dMD is a multithreaded tool able to utilize multiple CPU cores to increase the efficiency of descriptor calculations. The inputs to PyL3dMD are two mandatory inputs (LAMMPS input data and output trajectory files) and one optional input (number of cores for multiprocessing) for calculating 1982 3D descriptors. PyL3dMD requires minimal user intervention but can also be easily expanded to include more descriptors. The package can be used on all major platforms, including Windows, Linux, and macOS, by installing Anaconda. In addition, the PyL3dMD package can be easily integrated into post-processing scripts.

Future enhancements would include making it compatible with all atom styles, adding more and more complex descriptors, improving computational efficiency, generalizing file reading to make it compatible with other MD simulation tools, as well as integrating ML algorithms for QSPR modeling using the calculated descriptors.

This tool will enable scientific researchers to calculate a wide range of advanced 3D descriptors to quantify molecular chemistry and structure, and ultimately guide the design of advanced materials.

5.5. Summary of Supplementary Materials

The following supplementary materials can be found on GitHub at https://github.com/panwarp/PyL3dMD or in the `supplementary.zip` file:

- Sample script (Files: `Sample.py` and `Manual.pdf`)

- Sample usage example (File: `Manual.pdf`)
- Sample LAMMPS input data and output trajectory files (Folder: `LAMMPS Files`)
- A detailed list of descriptors calculated by PyL3dMD (File: `Manual.pdf`)
- Equations used to calculate descriptors (File: `Manual.pdf`)
- Installation instructions (File: `Manual.pdf`)

Chapter 6. Temperature-Dependent Density and Viscosity Prediction for Hydrocarbons: Machine Learning and Molecular Dynamics Simulations

6.1. Introduction

Lubricants are formulated by blending base oils with additives to enhance properties and meet performance specifications. A lubricant can contain up to 98% base oil, which is manufactured from crude oil or biological sources, or produced synthetically [8,232]. Most base oils are complex hydrocarbons which include paraffins, isoparaffins, aromatic, and naphthenic molecules of varying carbon number [233]. The molecular composition of a base oil directly affects its properties and, therefore, the performance of a lubricant [233]. Two of the most important properties of a base oil are viscosity and density because they are inherently linked to the lubricant hydrodynamics [124]. Viscosity and density directly affect friction and wear of a lubricated interface, and, in turn, the life of moving mechanical systems [10,234]. As such, it is critical to select an appropriate base oil when formulating lubricants with viscosity and density that meet the needs of a given application [233,234]. Although many experimental studies have been carried out on the physical, chemical, and thermodynamic properties of hydrocarbons, knowledge of such properties is not available for compounds that have not been characterized or synthesized yet. In addition, experimental measurements are not always possible for molecules generated *in silico* [208,235] because they have to be synthesized first which can be expensive and time-consuming. However, these limitations can be overcome by developing quantitative relationships between the molecular features of hydrocarbons and their density and viscosity at application conditions.

A recent study reviewed the advantages and limitations of several molecular modeling methods commonly used to predict density and viscosity [236]. Four molecular modeling methods were recommended: group-contributions (GC), equations of states (EoS), conductor like screening model for real solvents (COSMO-RS), quantitative structure-property relationships (QSPR), and molecular dynamics (MD) simulations. Particularly, QSPR, which is a molecular descriptor-based modeling, has been of significant interest in engineering as a means of predicting properties of complex materials or design of novel materials with improved properties. The QSPR method correlates measured physical or chemical properties with molecular features called descriptors [237–241]. QSPR has been widely used to predict biological and toxicological endpoints [206], especially in the pharmaceutical industry

[204,205]. Increasingly, it has also been used for predicting physicochemical properties [207–209].

In the past, many multivariate statistical tools, such as multiple linear regression (MLR), polynomial regression (PR), cluster analysis, principal component analysis (PCA), and partial least-squares regression (PLS), have been developed and applied to develop QSPR models [242–249]. For example, comprehensive descriptors for structural and statistical analysis (CODESSA) software was used to develop a five-descriptor MLR model for predicting dynamic viscosity of 337 organic molecules [242]. The descriptors included molecular weight, Randic connectivity index of order 3, charged surface area, maximum electrophilic reactivity index for a carbon atom, and maximum atomic orbital electronic population. The model predicted viscosities with a coefficient of determination (R^2) of 0.85. Later, the same software was used to build a five-descriptor QSPR model for estimating the viscosity of 360 organic molecules containing C, H, N, O, S, and/or halogens, with viscosities ranging from 0.197 to 1490 mPa·s [246]. This model had similar accuracy to that of the previously version [242]. Later, an improved five-descriptor PR model [247] was proposed to estimate the viscosities of the same 360 organic molecules. The R^2 for estimation was 0.91 [247]. This study identified mass, size, shape, and especially H-bonding as the most important predictors of viscosity [247]. Another study using CODESSA software developed a 16-variable 4-component PLS-based viscosity model for a small set of 46 molecules [248]. A study developed both MLR- and PLS-based viscosity models using 237 molecules of varying chemical structures with viscosities ranging from 0.197 to 19.9 mPa·s with an R^2 of 0.87 [243]. A year later, this study was extended and nine-descriptor viscosity models were proposed [244]. These models were developed using MLR and a simple neural network with a single hidden layer. Linear and nonlinear models were able to estimate viscosity of 124 test molecules with R^2 of 0.92 and 0.93 for the linear and nonlinear models, respectively [244]. The models in these studies [243,244] used experimentally determined values as descriptors, such as molar refraction, dipole moment, critical temperature, molar magnetic susceptibility, and cohesive energy. However, this approach is a limitation for discovery of new materials and the prediction of their properties.

Advances in machine learning (ML) had been pivotal in the development of more powerful QSPR models for accurately estimating the properties of lubricants [244,245,249–253]. In early work, an artificial neural network (ANN) was developed to predict the viscosity of 237 organic molecules of varying chemical structures with viscosities ranging from 0.197 to 19.9 mPa·s [253]. Another early study developed ANN models to predict the temperature dependence of viscosity of 440 molecules containing C, H, N, O, S, and halogens

at temperatures ranging from −120 to 160 °C [245]. This ANN model predicted viscosity of 79 test molecules with an R^2 of 0.88 [245]. In the same year, eight molecular descriptor-based models were developed to estimate the viscosity of 213 organic compounds [249]. The best model, based on neural networks, had R^2 equal to 0.86, an improvement over the estimations provided by the related MLR model [249]. Later, a study developed three types of models using group contributions [254], COSMO-RS σ-moment-based neural networks [255], and graph machines [256] approaches to estimate the viscosity of 300 pure liquids at 25°C for the cosmetic oils application [252]. A comparison of the viscosities of liquids of an independent set of 22 cosmetic oils shows that the graph machine approach provided the most accurate results with R^2 equal to 0.89 for 22 test molecules. More recently, a QSPR study developed a general feed-forward neural network to estimate dynamic viscosity of lubricant solvents and polymer-enhanced lubricants with 14 molecular descriptors [251]. The model was developed and evaluated using 1471 experimental viscosity values for either solvent or binary solvent mixtures and 52 experimental viscosity values for solvent/polymer mixtures of varying temperature from 10 to 100 °C. The model was reasonably accurate in predicting the dynamic viscosity of a test data set with an R^2 of 0.980 and 0.963 for lubricant solvents and solvent/polymer mixtures, respectively. A systematic model interpretation found the eccentricity index, Crippen MR, and Petitjean Number ($PetitjeanNumber$) [257] to be important predictors of viscosity.

QSPR models developed using advanced ML algorithm are more robust than experimental and MD simulation methods if suitable a set of molecular descriptors is used to develop the models. Generally, molecular descriptors can be of two types: static descriptors and dynamic descriptors, where static descriptors are determined directly from the chemical formulas of molecules, while dynamic descriptors are dependent on atom coordinates and, therefore, also on operating conditions. Although dynamic descriptors are more computationally expensive to obtain than static descriptors, they are more robust in predicting material properties [211,258–260]. However, there are no previous studies [43,61,74,77,78,231,261,262] on QSPR modeling of viscosity and density using dynamic descriptors obtained from MD simulations. In addition, static descriptors and dynamic descriptors have not been used together in QSPR models and compared side-by-side.

In this study, we developed temperature-dependent viscosity and density models for 305 complex hydrocarbons based on a QSPR approach using both static and dynamic molecular descriptors. We developed simple and easily interpretable ML models using Gaussian process regression (GPR). Various parameters were used to assess the quality of models during training, validation, and testing. The predictive capabilities of these models were tested

by applying the models to subsets of molecules excluded from the training and validation data sets. Finally, to determine the impact of each descriptor on the model predictions, detailed model-agnostic interpretation was conducted using various ML interpretation tools. The model interpretation and significance of model terms may assist in selection of existing hydrocarbons or design of new molecules for base oils with desired viscosity and density. The model development approach presented in this study can also be extended to prediction of other important properties of materials.

6.2. Methods and Materials

This section outlines the design and training process of the descriptor-based ML models. The flowchart in Figure 6.1 shows the overall workflow of the ML approach. Step 1 is the collection of experimental data used to train, validate, and test the models. Step 2 is calculation of a large set of molecular descriptors or model predictors (predictors = descriptors + operating conditions). Step 3 is selection of the significant predictors from a large set of predictors using the least absolute shrinkage and selection operator (LASSO) regularization, F^1 test, and elimination of the strongly correlated predictors using correlation/collinearity analyses. Step 4 is development of models with all possible combination of significant predictors using GPR. Step 5 is the optimization of hyperparameters and then selection of the best models using model assessment parameters: coefficient of determination (R^2), root mean squared error (RMSE), and the variance inflation factor (VIF) values. Step 6 is model-agnostic interpretation of the best models by partial dependency plots (PDP), individual conditional expectation (ICE) plots, average local-interpretable model-agnostic explanations (LIME), and relative decrease in R^2 values due to trimming a predictor. Lastly, Step 7 is evaluation of the final models using test data.

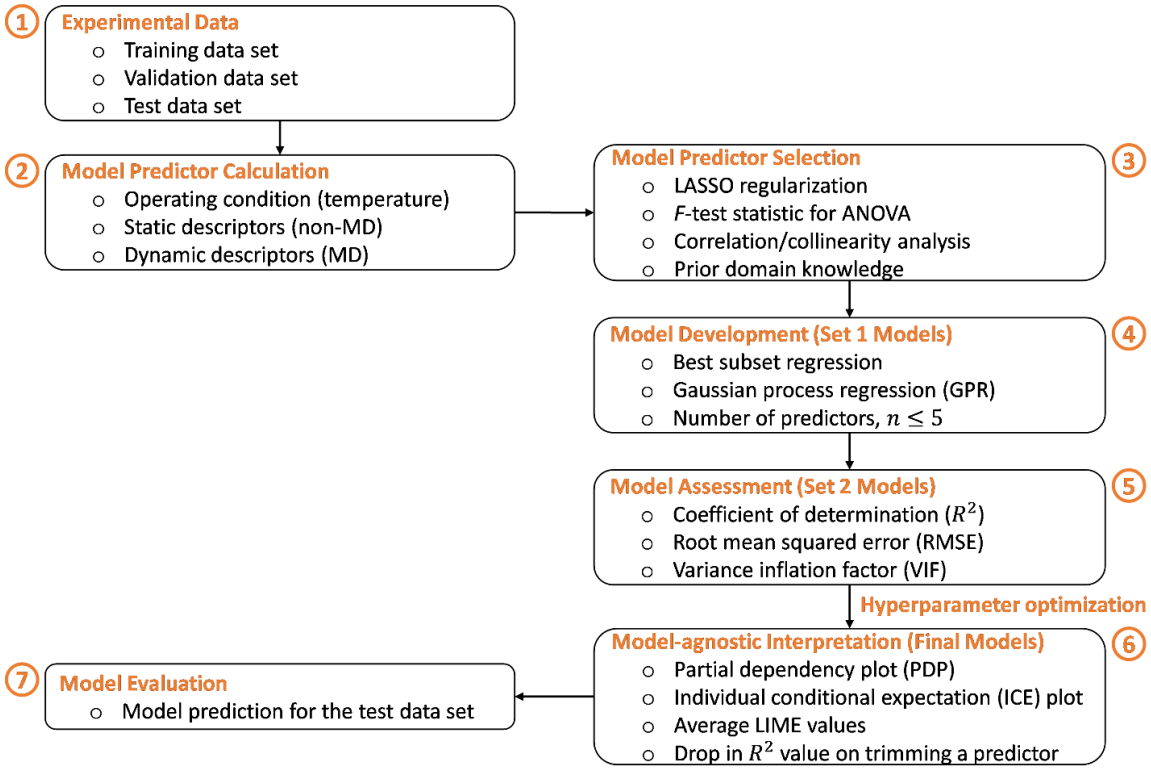


Figure 6.1 The overall workflow of the ML approach to design, train, and evaluate the models.

6.2.1. Experimental Data

Dynamic viscosity and density of 305 pure hydrocarbons (C_8 to C_{50}) at a wide range of temperatures were obtained from the American Petroleum Institute (API) Research Project 42 [263]. These 305 hydrocarbons include n-paraffins, branched-paraffins, 1-olefin, branched-olefins, non-fused ring naphthene, fused ring naphthene, non-fused ring aromatic, and fused ring aromatics. Structures of some of the hydrocarbons are shown in Figure 6.2 to illustrate the diversity of molecule structures. The molecular weights of the 305 hydrocarbons range from 110.20 to 703.30 g/mol. The viscosities and densities range from 0.29 cP to 2.00×10^4 cP and 0.67 g/cc to 1.12 g/cc, respectively. The viscosities and densities of these molecules with their molecular formulas and simplified molecular input line entry system (SMILES) [229] codes can be found in the supplemental document (Tables C.1 and C.2). In addition, schematics of all molecules can be found in the supplementary zip file (schematics.zip). Viscosity and density data at atmospheric pressure and temperatures ranging from 0 °C to 135 °C were used to develop the models. The API Research Report provided both density and viscosity for most hydrocarbons. However, for some hydrocarbons or temperatures, only density

or only viscosity was reported. In total, 1292 viscosity data points and 1474 density data points were included in the model development.

Due to the large data set, we used the holdout cross-validation technique. Data were divided randomly into three partitions to develop and assess the models: training, validation, and test data sets. First, 70% of the 305 molecules (215 molecules) were used to train the models, and then 15% of the 305 molecules (45 molecules) were used to validate the models during development. Lastly, the remaining 15% of the 305 molecules (45 molecules) that were not in the training and validation data sets were selected to assess the accuracy of the predictions from the developed models. All molecules are listed in Table C.1, each identified with a partition ID, either 1, 2, or 3, to indicate if it was in the training, validation, or test data set. All the experimental data can be found in the Appendix C.

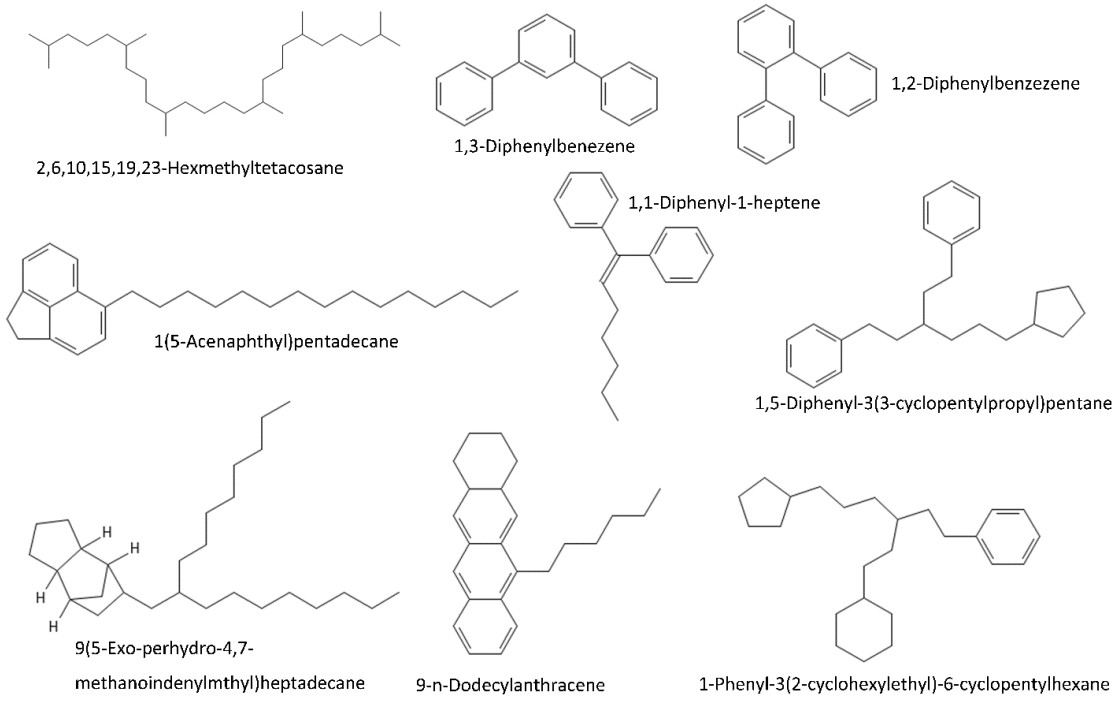


Figure 6.2: Structures of representative hydrocarbons used in the model training, validation, or test data sets.

6.2.2. Model Predictors

Model predictors were temperature and molecular descriptors. Temperature was included as an operating condition predictor since both viscosity and density are inversely related to temperature [21–24]. Molecular descriptors are mathematical representations of the physical and chemical nature of molecules [259]. In this study, we classified molecular descriptors into two categories

based on the level of molecular representation required for calculating them, either static or dynamic. Static molecular descriptors are one-dimensional (1D) and two-dimensional (2D) descriptors that do not require three-dimensional (3D) coordinates of the atoms in a molecule for their calculation. Dynamic molecular descriptors are 3D descriptors, commonly known as geometric descriptors, which require the 3D coordinates of atoms for their calculation. Dynamic descriptors are more robust and better able to capture the conformations of molecules at different operating conditions such as temperature, pressure, and speed. However, dynamic descriptors require a higher computational cost to calculate. Since all predictors except temperature were molecular descriptors, the terms "predictors" and "descriptors" can be effectively used interchangeably.

A total of 1444 static descriptors were obtained using an open-source software, PaDEL [212], by providing SMILES codes [229] of the molecules in Table C.1. The details of all 1444 static descriptors can be found in the supplementary information. In addition, 156 dynamic descriptors were determined using MD simulations. Of these, 57 dynamic descriptors were directly obtained from the simulations and 99 dynamic descriptors were calculated by postprocessing the atomic trajectories from the simulations. The 57 dynamic descriptors directly from MD such as stress tensor, energies, density, volume, and dipole moment. The rest of the dynamic descriptors were calculated via our open-source python-package PyL3dMD [262] using the 3D coordinates, connectivity, and charge of atoms obtained at each timestep of the MD simulation. Since it is not possible to provide the physical and chemical significance of all descriptors used in this study, the significance of only highly correlated predictors is provided with the model interpretation. The definition of each descriptor used in this study can be found in the supplementary information.

For each molecule, a cubic model system with periodic boundaries containing around 5000 atoms and 5.0 nm^3 was created using an open-source software, Packmol [264]. The model systems had volumes of around 5.0 nm^3 to minimize finite-size effects across the periodic boundary [169] and to reduce pressure and stress fluctuations [171], enabling accurate and reliable simulations. All atomic interactions were described using the All Atom Optimized Potentials for Liquid Simulations (OPLS) [122], one of the most popular and accurate potentials for calculating transport properties of hydrocarbons. The OPLS parameters for the hydrocarbons were obtained using the LigParGen [265] and BOSS [266] open-source software packages. We used the CM1A-LBCC [267] charge model to assign charges to the atoms of the hydrocarbons. Finally, dynamic simulations were run with a time step of 1.0 fs using Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software [121].

A previously developed, robust equilibration molecular dynamics (EMD) simulation approach was followed to optimize the geometries of molecules at atmospheric pressure and temperatures consistent to experimental temperatures [159]. First, energy minimization of the system was performed using the conjugate gradient algorithm. Second, the system was heated to 1000 K for 0.25 ns in the canonical (NVT) ensemble to achieve homogeneity. Third, the system density was equilibrated at 1.0 atm and a target temperature for 1.0 ns in the isothermal–isobaric (NPT) ensemble using the Nosé–Hoover thermostat and barostat [125,126], with damping coefficients of 100 and 25 fs, respectively. Fourth, while maintaining a constant temperature for 0.5 ns in the NVT ensemble, the simulation box was deformed until the density of the fluid reached the average density computed from the previous NPT simulations. Finally, the system was equilibrated using the final configuration from NVT as the initial configuration for 0.25 ns in the microcanonical (NVE) ensemble. Following this, two independent production runs of 1.0 ns and 3.0 ns were carried out in the NPT and NVE ensembles, respectively. To develop the density models, dynamic descriptors were calculated from the NPT ensemble, since models that allow volume fluctuation are commonly used to calculate density in MD. For the viscosity models, dynamic descriptors were calculated from the NVE ensemble, to be consistent with typical simulation methods used to calculate viscosity from MD simulations with the Green-Kubo formula [78,231]. The only exception is *rho*, density calculated from NPT simulations, is used as a descriptor to develop both density and viscosity ML models. However, a key difference between directly calculating density or viscosity from MD simulations and using simulations in the same ensembles to calculate descriptors is that the latter requires significantly less computational time. For example, here, we used only 3 ns of NVE simulation data to predict viscosity using molecular descriptors from the simulation whereas, in our previous studies [78,231], we had to run simulations for 400 ns to accurately calculate the viscosity of a lubricant. The 3D coordinates of the atoms and directly computed dynamic descriptors were stored every 1000 fs of each production run. The dynamic descriptors were averaged over the second half of production simulations. The LAMMPS scripts and data files of all molecules and the calculated static and dynamic descriptors are provided in the supplementary information.

6.2.3. Model Predictor Selection

Developing models of all possible combinations of a large set of predictors is inefficient and only feasible using supercomputing resources. Predictor selection reduces the dimensionality of data by selecting only a subset of predictors to create a model that accurately predicts measured responses. The

primary objectives of predictor selection techniques are to improve prediction performance, provide faster and more cost-effective predictors, and improve model interpretability [268]. Therefore, after gathering a large complex set of potential predictors, the next step was to remove redundant, unimportant, and strongly correlated predictors to avoid unreliable and unstable estimates from the regression models. Therefore, LASSO regularization [269,270], F-test [271], correlation/collinearity analysis [272,273], and prior domain knowledge were used to remove redundant and strongly correlated predictors and select the most important predictors of the response variable (density/viscosity).

Here, we explain the predictor selection approach for static descriptors used to develop viscosity models. We started with 1444 static descriptors. After removing descriptors with any missing values, infinite values, or the same values for all molecules, we were left with 944 static descriptors. Next, a LASSO fit with 10-fold cross-validation was performed, and descriptors in the sparsest model within one standard error of the minimum mean squared error (MSE) (*LambdaMinMSE*, as shown in Figure 6.3a) were selected. In LASSO regularization, the coefficients of covariates that were strongly correlated with one another or are less relevant to the response variable (in this case viscosity) were eliminated from the pool of predictors. After performing LASSO regularization, we were left with only 37 identified important predictors. Second, using an F-Test, we ranked the importance of all 944 static descriptors for the response variable, as shown in Figure 6.3b. The F-Test is the statistic used for analysis of variance (ANOVA) to examine the importance of each predictor individually. p-value, also known as probability value, is a statistical measurement used to validate a hypothesis against observed data. A small p-value or a large negative $log(p)$ value of the test statistic indicates that the corresponding predictor is important. Here, the negative logarithmic p-value is the predictor's score, which indicates the predictor's importance to the response variable. Using scores of F-tests with prior domain knowledge of important predictors, we included highly important predictors that LASSO regularization might have removed from the pool. Third, a correlation matrix was used to assess the cross-correlation of the predictors and remove strongly correlated predictors, as shown in Figure 6.3c. The correlation matrix is a standard measure of the strength of pairwise linear relationships. Finally, predictors not eliminated at this modeling step were considered the most significant and used for developing viscosity models. The same predictor selection approach was used for dynamic descriptors and for the density models.

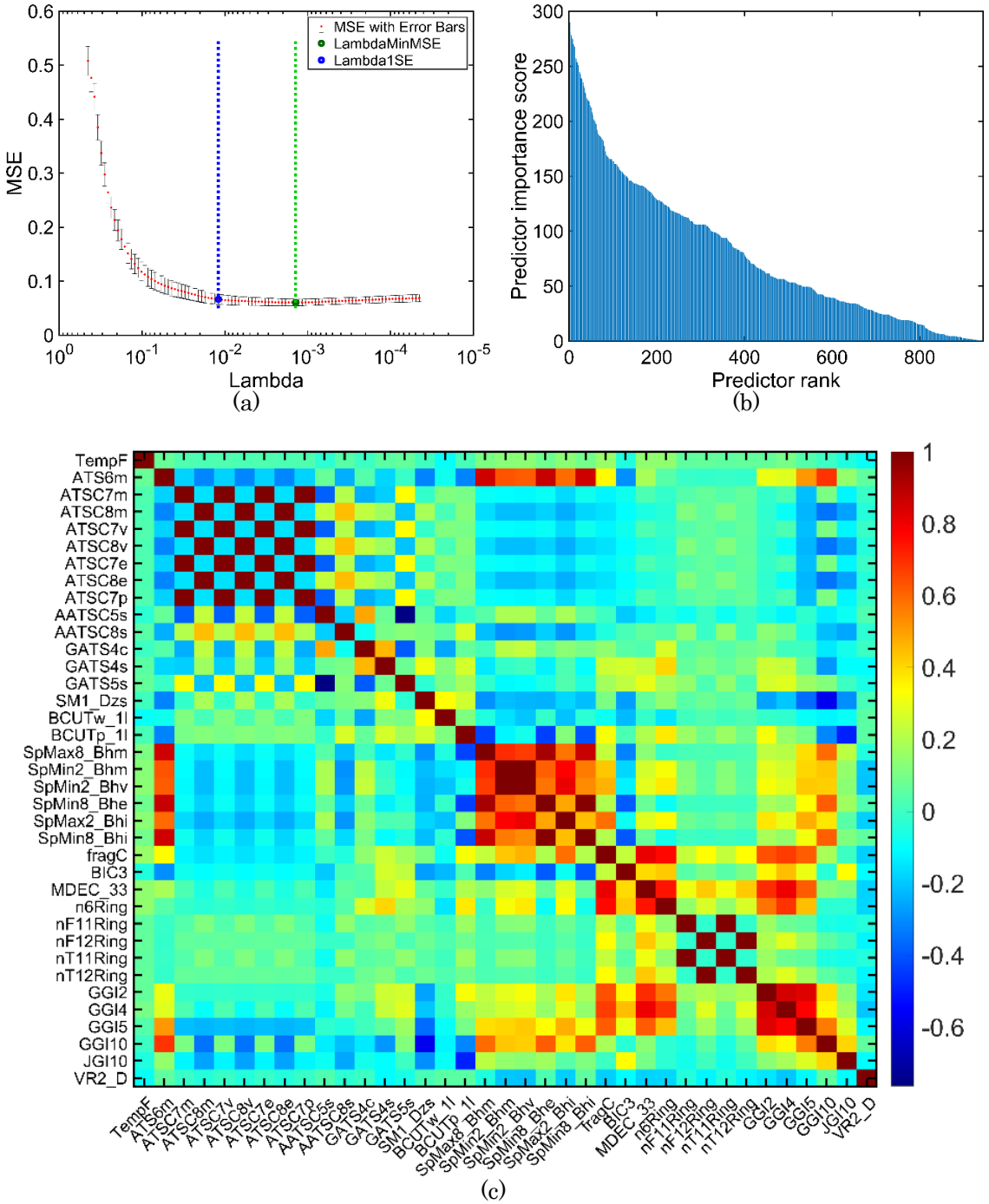


Figure 6.3: (a) MSE of the LASSO fit using 10-fold cross-validation. The lambda that results in the lowest MSE is the green dotted line whereas the blue dotted line is the lambda that is within one standard error of the lowest MSE. (b) Score of all predictors using the F-test where, in this example, the most important predictor is given rank 1 and the least important predictor is given rank 944. (c) Pairwise linear correlation coefficients of the descriptors where the dark

red and dark blue represent highly positively and negatively correlated predictors, respectively.

6.2.4. Model Development and Assessment

After selecting important predictors and randomly dividing the experimental data into three data sets for training, validation, and testing, we developed ML algorithm-based models. GPR was used because of its flexibility and tractability. GPR models are nonparametric kernel-based probabilistic models [274]. GPR was combined with the best subset regression approach to develop GPR-based models for each possible combination of predictors from the pool of selected important predictors. In this approach, all possible models were developed with up to five predictors or a significant increase of the R^2 [275] when increasing the number of predictors. The models generated at this stage were defined as the first set of models. The workflow of the model development step is shown in Figure 6.4.

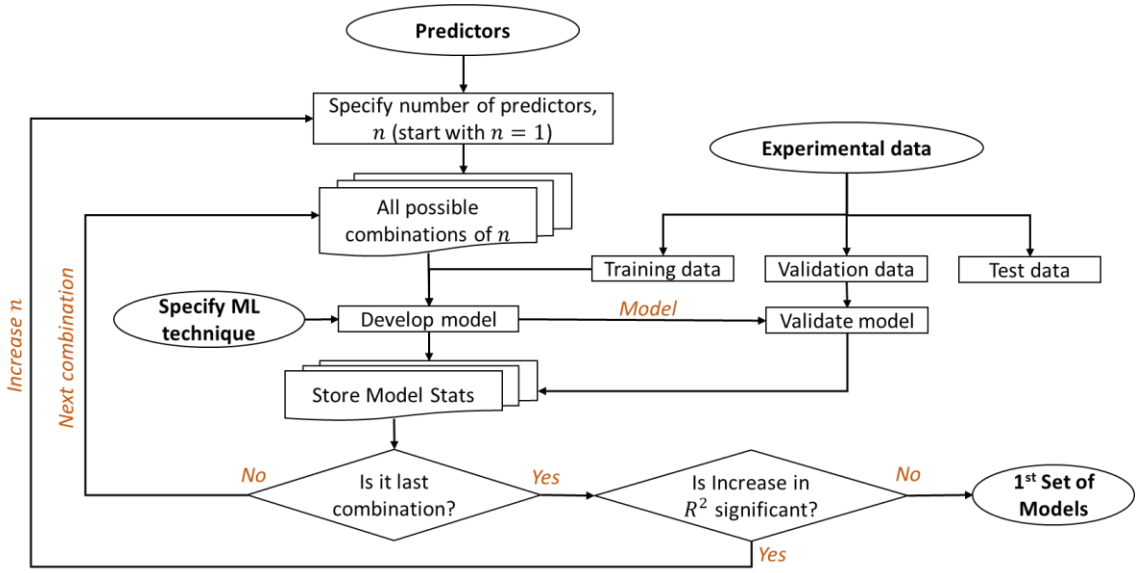


Figure 6.4: Workflow of the model development step.

In this step of model development, a holdout cross-validation technique was implemented to validate trained models and the statistics of training and validation were stored to assess the quality of the models. These statistics include R^2 and RMSE values for the training and validation data sets and VIF of each predictor. The predictability of models was assessed based on the R^2 and RMSE values. The R^2 is a statistical measure of fit that quantify variation of the response variable that is predictable by the predictor(s) in a regression model. The R^2 was calculated using Equation (6.1).

$$R^2 = \frac{\sum_{i=1}^{N}\left(y_i^{exp} - \bar{y}\right)^2 - \sum_{i=1}^{N}\left(y_i^{exp} - y_i^{pred}\right)^2}{\sum_{i=1}^{N}\left(y_i^{exp} - \bar{y}\right)^2} \quad (6.1)$$

Here, N is the total number of data points, i is the i^{th} data point, y_i^{exp} is the experimental value of the response variable, y_i^{pred} is the model predicted value of the response variable, and $\bar{y}$ is the mean experimental value of the response variable. The RMSE, which is a measure of the difference between values predicted by a model and values observed by experiment, was calculated using Equation (6.2).

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i^{exp} - y_i^{pred}\right)^2} \quad (6.2)$$

VIF was used to evaluate the multicollinearity between three or more predictors with the response variable and with each other. VIF was calculated using Equation (6.3) [272].

$$\text{VIF}_j = \frac{1}{1 - R_j^2} = diag(\mathbf{C}^{-1}) \quad (6.3)$$

Here, $\mathbf{C}$ is the correlation matrix or matrix of the correlation coefficient, and R_j^2 is the R^2 of predictor j on the remaining predictors. When the variation/trend of a predictor j is nearly a linear combination of the other predictors, then R_j^2 is close to 1 and the *VIF* for that predictor is correspondingly large. If R_j^2 is 0 (no collinearity), then VIF is 1, which is the lowest possible value of VIF. We used VIF > 5 as a benchmark for the presence of multicollinearity [276] and discarded models with VIFs higher than 5.

The 100 models with the highest R^2 value, lowest RMSE value, and VIF values for all predictors less than or equal to 5 were selected as the 2nd set of the models. The predictive performance of these models is sensitive to the numerous hyperparameters in the GPR algorithm. Therefore, hyperparameters of the models in the 2nd set were optimized using a 5-fold cross-validation technique. The multidimensional combinatorial hyperparameter space was explored using the Bayesian optimization [277,278] algorithm to tune hyperparameters. Bayesian optimization was chosen because, unlike other optimization techniques, it uses information from past evaluations of the function and does not only depend on local gradient and

Hessian approximations of function [279]. This allows the optimization search to reach the minimum rapidly, even for nonconvex functions [280].

After retraining the top 100 models with optimized hyperparameters, we chose the best model for each response variable as the model with the highest R^2 value, the lowest RMSE value, VIFs ≤ 5, and for which the trends of PDPs [281,282] were consistent with the expected physical behavior. For example, it is commonly known that the viscosity and density of liquids decrease as temperature increases. If a model does not match the expected trend, the model is not correct, even if it has a perfect R^2 of 1.0 (or 100%). PDP depicts the marginal effect of a predictor on the outcome of a model, and the extent of change in the response variable to a change in a predictor indicates the global importance of that predictor. In addition, for two models with similar statistics, the model with simple (easy to understand and calculate) descriptors was preferred.

6.2.5. Model Interpretation and Evaluation

To better understand the predictions, we systematically interpreted our best models and their predictors. We conducted model-agnostic interpretation using PDPs [281,282], ICE [282], and LIME [283]. PDP is a tool to investigate global importance (the contribution of a predictor to the overall prediction of data) whereas ICE and LIME are tools used to investigate the local importance (the contribution of a predictor to the prediction of each data). The PDPs do not reveal hidden dependencies because they only show averaged relationships between a predictor and response variable. ICE plots can be used to find interactions among model variables and unusual subgroups in the data sets [284]. Therefore, to investigate heterogeneities of partial dependence originating from different observations, we generated ICE plots for each predictor of the best models. As the name suggests, the LIME value or weight shows local importance, however, the mean of the LIME values for all data points can be used as a global representation of predictor importance. A positive (negative) mean LIME value implies a positive (negative) relationship of the predictor with respect to the response variable. The importance of predictors was analyzed by trimming predictors one at time from the best model and observing the performance (R^2) of the trimmed model. Finally, the best models for viscosity and density were evaluated with a new set of hydrocarbons (test data set) to the verify their predictive performance over a wide range of temperatures.

6.3. Results and Discussion

6.3.1. Density Models

We were able to achieve good temperature-dependent density models with three or fewer predictors. The three best static descriptor-based models for temperature-dependent density with one, two, and three static predictors are Equations (6.4), (6.5), and (6.6), called Model I, II, and III. Model I is only a function of Broto-Moreau autocorrelation-lag 2/weighted by Sanderson electronegativities ($AATS2e$). $AATS2e$ is a spatial autocorrelation calculated from molecular graph, that is connectivity of atoms of a molecule [285] where e in $AATS2e$ is the Sanderson electronegativity [286] of atoms in a molecule, whereas 2 in $AATS2e$ is the lag or the topical distance between two connected atoms in a molecule. Therefore, it is a measure of molecular connectivity and complexity. Example calculation of these static descriptors can be found elsewhere [259]. Model II is a function of temperature (T) in °F and $AATS2e$. Model III is a function of temperature, conventional bond order ID number of order 3 ($piPC3$), and fraction of rotatable bonds, including terminal bonds ($RotBtFrac$). Conventional bond order ID number is a molecular weighted path number calculated from weighting graph edges (bonds) with conventional bond order, which is defined as being equal to 1, 2, 3, and 1.5 for single, double, triple, and aromatic bonds, respectively. $piPC3$ is a conventional bond order weighted measure of molecular connectivity and complexity [259]. The conventional bond orders for single, double, triple, and aromatic bonds are 1, 2, 3, and 1.5 [259]. $RotBtFrac$ is the fraction of rotatable bonds over the total number of bonds in a molecule [259]. Rotatable bonds are bonds that meet the three following criteria: (a) single bond connected by heavy atoms with the heavy atoms connected to at least one atom (including hydrogen atom), (b) the external bond by which the heavy atom is connected must not a triple bond unless the triple bonded atom is connected to another atom, and (c) the bond must not be part of a ring [259].

Model I: $$\rho(T) = f(AATS2e) \tag{6.4}$$

Model II: $$\rho(T) = f(T, AATS2e) \tag{6.5}$$

Model III: $$\rho(T) = f(T, piPC3, RotBtFrac) \tag{6.6}$$

The best three dynamic descriptor-based models for density with one, two, and three dynamic predictors are Equations (6.7), (6.8), and (6.9). Model I is only a function of simulation-calculated density (rho) from the NPT ensemble. Model

II is a function of rho and the radius of gyration (R_g) of the molecule which quantifies molecular size. Model III is a function of rho, R_g, and energy due to van der Waals interactions ($evdwl$). Note that rho is different from ρ, although both are density: rho is a dynamic descriptor calculated from the simulations and ρ is the experimentally measured fluid density predicted by the ML model. We include rho as a descriptor because it can be calculated from a very short simulation and is part of the PyL3dMD python package [262]. The density models with dynamic descriptors excluding the simulation-calculated density are also provided in the supplementary information.

Model I: $$\rho(T) = f(rho) \tag{6.7}$$

Model II: $$\rho(T) = f(rho, R_g) \tag{6.8}$$

Model III: $$\rho(T) = f(rho, R_g, evdwl) \tag{6.9}$$

Table 6.1 Model assessment parameters for the density models with static and dynamic descriptors.

Parameter		Static Descriptors			Dynamic Descriptors		
		Model I	Model II	Model III	Model I	Model II	Model III
Training	R^2	0.893	0.989	0.999	0.983	0.994	1.000
	RMSE	0.025	0.008	0.003	0.010	0.006	0.001
Validation	R^2	0.910	0.993	1.000	0.985	0.994	1.000
	RMSE	0.024	0.007	0.001	0.010	0.006	0.000
Test	R^2	0.874	0.935	0.981	0.971	0.977	0.988
	RMSE	0.029	0.021	0.011	0.014	0.012	0.009
Average R^2		0.892	0.972	0.993	0.980	0.988	0.996
Maximum VIF		1.000	1.003	1.316	1.000	1.203	4.113

Table 6.1 lists the R^2 and RMSE values for the density models with static and dynamic descriptors for the training, validation, and test data sets. The average R^2 values for the training, validation, and test data sets are also reported to enable comparison of the models. The maximum VIF for each predictor is also given to indicate the degree of multicollinearity. Due to the single predictor in Model I with either static and dynamic descriptors, the VIF value is a perfect 1, but the VIF increased as the number of predictors increased in Models II and III. The statistics in Table 6.1 show that the model (Equation (6.4)) with a single static descriptor $AATS2e$ was able to reach an accuracy of 89.2% in predicting the density of the hydrocarbons. When a temperature term was added to the model, that is, in Model II (Equation (6.5)),

the R^2 for the test data set increased to 97.2%. The best model with the static descriptors, Model III (Equation (6.6)), has R^2 values of 99.9%, 100.0%, and 98.1% for the training, validation, and test data sets.

From the statistics in Table 6.1, the density models with a single dynamic descriptor (Equation (6.7)), *rho* calculated from only 1.0 ns of simulation time, was able to achieve an accuracy of 98.0% in predicting density. This is higher than any single static descriptor and any combination of two static descriptors. When other simulation-calculated descriptors were included, i.e., Models II and III in Equations (6.8) and (6.9), the R^2 values increased to around 99%. The best density model with dynamic descriptors, Model III (Equation (6.9)), had R^2 values of 100.0%, 100.0%, and 98.8% for the training, validation, and test data sets.

Figures 5.5(a) and 5.5(b) show experimental density and density predicted by the best model (Model III) with static and dynamic descriptors for all data sets over a wide temperature range. The blue dashed lines represent ideal predictions. The model-predicted density for the training, validation, and test data sets is shown as black circles, red squares, and green triangles, respectively. From the statistics in Table 6.1 and Figure 6.5, we can see both models performed exceptionally well with only three descriptors. Furthermore, the model with three dynamic descriptors performed slightly better than the model with three static descriptors.

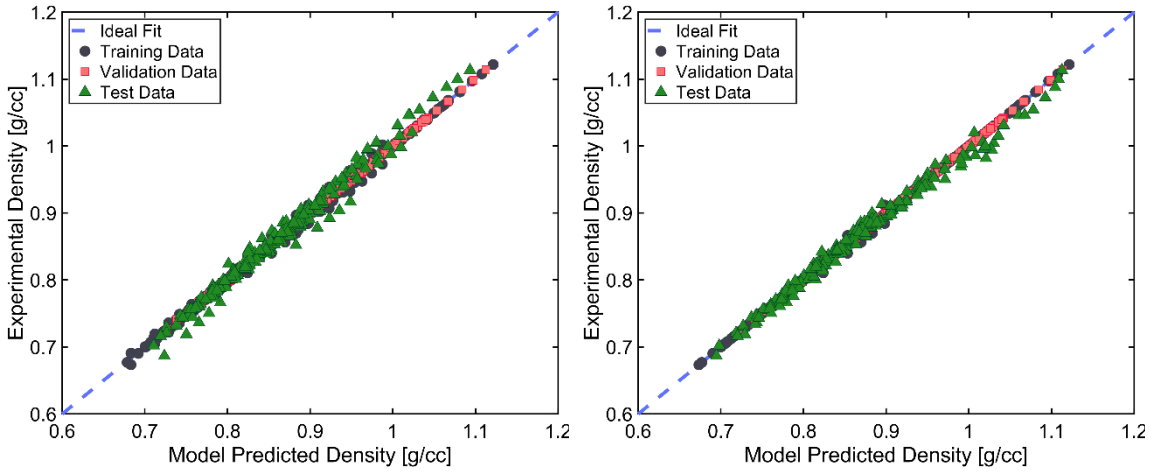


Figure 6.5: Experimental density vs. density predicted by the best models (Model III) with (a) static and (b) dynamic descriptors for the training (black circles), validation (red squares), and test data (green triangles) sets. The blue dashed lines represent the ideal prediction.

We can visualize the relationships between each model predictor in a trained regression model and model-predicted responses using the PDPs and ICE

plots. In Figure 6.6, the circle symbols show the predicted response for each data point. The PDP (red line) shows the averaged relationship, whereas ICE plots (gray lines) show an individual dependence for each observation [282], resulting in one line per observation. The PDPs are offset such that the y-axis starts at zero to illustrate cumulative effects and the importance of each predictor in the model. Figures 5.6(a) and 5.6(b) show the relationship between the model-predicted density and each static predictor in Equation (6.6) and each dynamic descriptor in Equation (6.9), respectively.

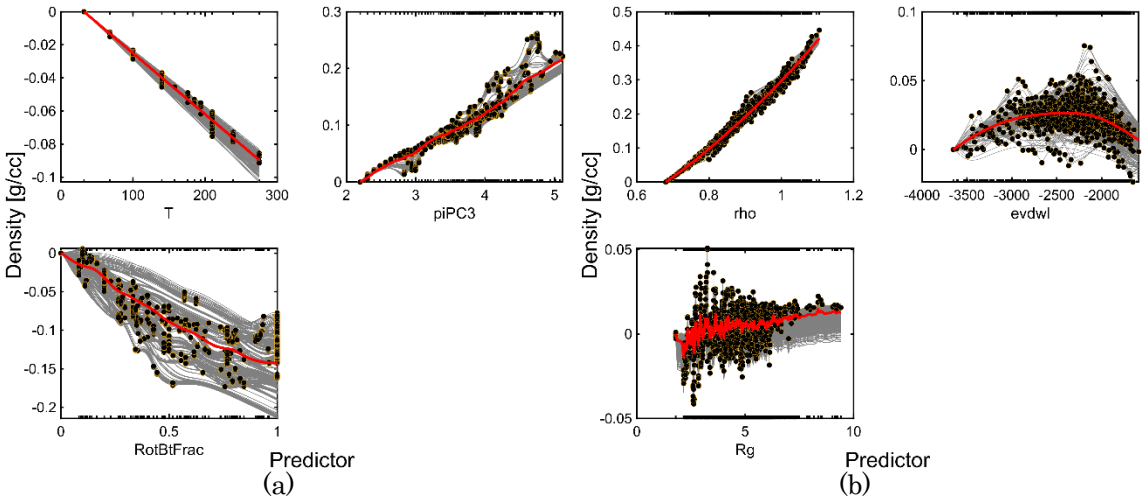


Figure 6.6: The partial dependence (red curve) and the individual conditional expectation (grey curves) of each predictor in the best density models (Model III) with (a) static and (b) dynamic descriptors. The scattered circular symbols represent the relationship between a predictor variable and density for each observation. The data are offset so that the density starts from zero to better to illustrate the cumulative effect of a predictor on density.

PDPs and ICE plots show that the density decreases with increasing temperature (T) in °F and fraction of rotational bonds in the molecules ($RotBtFrac$) but increases with increasing $piPC3$. A larger value of $RotBtFrac$ indicates greater ease of rotation of the backbone in the molecules (i.e., chain flexibility) [287]. In Figure 6.6(b), it must be noted that the negative sign of $evdwl$ suggests that the interaction is driven by an attractive force. Therefore, a higher negative value of $evdwl$ means stronger van der Waals interactions. Figure 6.6(b) shows that the experimental density and rho are linearly correlated, as expected. It also shows that the density increases with R_g (higher for a bigger molecule) and decreases with increasing van der Waals interactions energy ($evdwl$). The ICE plots clearly show that density decreases with increasing T, $RotBtFrac$, and $evdwl$, but increases with increasing $piPC3$ and R_g for most of the observations. Figures 5.7(a) and 5.7(b) show the average LIME values over all observations for the models with static and dynamic descriptors. The sign of the LIME values also indicates that density of

hydrocarbons decreases with increasing T, $RotBtFrac$, and $evdwl$, but increases with increasing $piPC3$ and R_g.

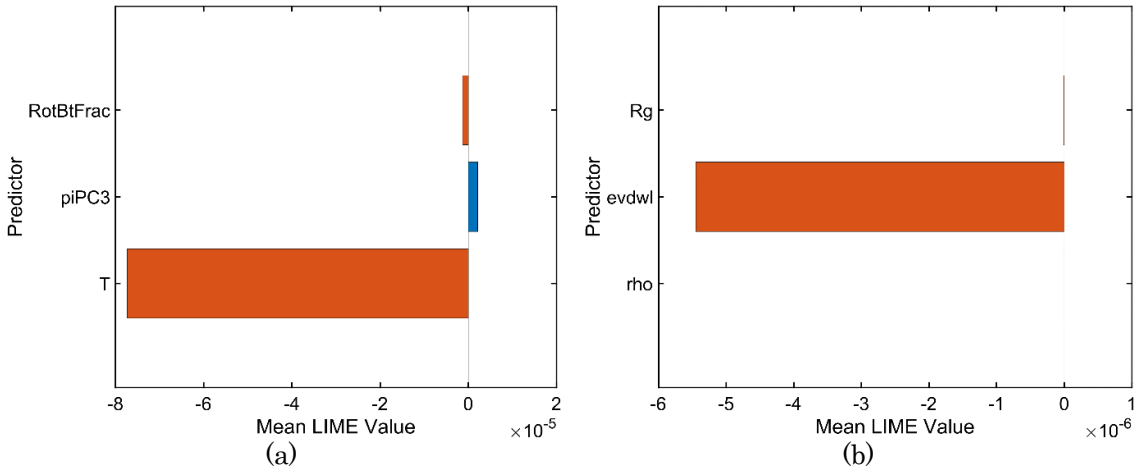


Figure 6.7: The average LIME value for each predictor in the best density models (Model III) with (a) static and (b) dynamic descriptors. The orange and blue colors represent negative and positive relationships between a predictor and the response variable. The size of a bar represents the overall importance of a predictor.

To rank the importance of each predictor in the models, we removed each predictor one at a time and observed the performance of the model with a dropped predictor. Table 6.2 gives the average R^2 of the training, validation, and test data sets for each predictor when it was trimmed from the best density model. For instance, when the T term was dropped from the best density model with static descriptors, the average R^2 value decreased from 99.3% to 90.0%, that is, by 9.3%. A larger decrease in R^2 indicates more importance. The decrease in performance of the best model was 9.3%, 18.3%, or 21.4% when the T, $RotBtFra$, or $piPC3$ term was dropped from the best density models. Therefore, the order of importance of the predictors in the best density model with static descriptors is $piPC3 > RotBtFrac > T$. Similarly, for the best density model with dynamic descriptors, the order of predictor importance is $rho > R_g > evdwl$. When $evdwl$ was dropped from the best density model with dynamic descriptors, the performance decreased by only 0.8% which suggests that this term could be permanently from the model. This means that the most important predictors of density are T, $RotBtFrac$, $piPC3$, R_g, and rho.

Table 6.2: Performance of the best density models when a predictor was removed from the models.

Static Descriptors (Equation (6.6))			**Dynamic Descriptors** (Equation (6.9))		
Term Dropped	Avg. R^2	Drop in Avg. R^2	Term Dropped	Avg. R^2	Drop in Avg. R^2
None	99.3%	0.0%	None	99.6%	0.0%

T	90.0%	9.3%	$evdwl$	98.8%	0.8%
$RotBtFrac$	81.0%	18.3%	R_g	98.1%	1.5%
$piPC3$	77.9%	21.4%	rho	82.0%	17.6%

We also developed density models with combined static and dynamic descriptors. The combined models performed slightly better (R^2 of 0.997) than the models with only static (R^2 of 0.993) and only dynamic descriptors (R^2 of 0.996). Therefore, they were not analyzed further but are provided in the supplementary information.

6.3.2. Viscosity Models

Instead of training models directly for viscosity (η), we trained the models by taking logarithmic viscosity ($\log\eta$), based on the knowledge the viscosity decreases exponentially with temperature. Unlike density, we were not able to achieve good temperature-dependent viscosity models with one or two predictors. The best three models with two, three, and four static predictors for temperature-dependent viscosity are Equations (6.10), (6.11), and (6.12). Model I is a function of temperature (T) in °F and the first kappa shape index ($Kier1$) [288]. Model II is a function of T, $RotBtFrac$, and the molecular weight (MW). Model III has the same terms as Model II plus the $PetitjeanNumber$ [257]. The terms $RotBtFrac$ in Model II and $RotBFrac$ in Model III are the fraction of rotatable bonds over the total number of bonds, including ($RotBtFrac$) and excluding ($RotBFrac$) terminal bonds, respectively. $Kier1$ is a connectivity descriptor which quantifies the complexity in connectivity of the molecule. $PetitjeanNumber$ is a topological anisometry descriptor which quantifies the molecular shape [257]. It is calculated from the generalized radius and diameter of the molecule [257]. Example calculations of these static descriptors can be found elsewhere [257,259].

Model I: $$\log\eta(T) = f(T, Kier1) \quad (6.10)$$

Model II: $$\log\eta(T) = f(T, RotBtFrac, MW) \quad (6.11)$$

Model III: $$\log\eta(T) = f(T, RotBFrac, MW, PetitjeanNumber) \quad (6.12)$$

The best three models for viscosity with two, three, and four dynamic predictors are Equations (6.13), (6.14), and (6.15). Model I is a function of rho and kinetic energy (ke). Model II is a function of rho, energy due to improper interactions ($eimp$), and acylindricity (c). $eimp$ quantifies the stiffness of the molecule, proportional to the inverse of rotatable bonds in the molecule. c is a

measure of cylindricity in the distribution of atoms in a molecule. c is zero when a molecule is cylindrically symmetric and increases as the molecule deviates from this shape [190]. Model III is a function of rho, $eimp$, T, and a diagonal component of the moment of inertia tensor (I, i.e., size or distribution of atomic mass from the center of mass of a molecule). It was found that any diagonal component (I_{xx}, I_{yy}, I_{zz}) of the moment of inertia tensor resulted in similar predicting performance when used in Equation (6.15), likely because the dynamic descriptors were calculated from equilibration MD simulations with a cubic simulation box.

Model I: $$\log\eta(T) = f(rho, ke) \quad (6.13)$$

Model II: $$\log\eta(T) = f(rho, eimp, c) \quad (6.14)$$

Model III: $$\log\eta(T) = f(T, rho, eimp, I_{zz}) \quad (6.15)$$

Table 6.3: Model assessment parameters for the dynamic viscosity models with static and dynamic descriptors.

Parameter		Static Descriptors			Dynamic Descriptors		
		Model I	Model II	Model III	Model I	Model II	Model III
Training	R^2	0.949	0.989	0.991	0.942	0.994	0.999
	RMSE	0.161	0.076	0.070	0.173	0.010	0.020
Validation	R^2	0.970	0.988	0.989	0.95	0.987	0.998
	RMSE	0.129	0.082	0.078	0.167	0.003	0.031
Test	R^2	0.834	0.936	0.952	0.809	0.876	0.932
	RMSE	0.275	0.171	0.148	0.324	0.293	0.176
Average R^2		0.918	0.971	0.977	0.900	0.952	0.977
Maximum VIF		1.007	1.229	1.409	1.021	1.115	1.874

Table 6.3 gives the R^2 and RMSE values of the viscosity models with static and dynamic descriptors for the training, validation, and test data sets. The statistics in Table 6.3 show that the R^2 values for the best model with static descriptors are 99.1%, 98.9%, and 95.2% for training, validation, and test data sets. Similarly, the R^2 values for the best model with dynamic descriptors are 99.9%, 99.8%, and 93.2% for training, validation, and test data sets. The average R^2 value of 97.7% is same for the models with static or dynamic descriptors.

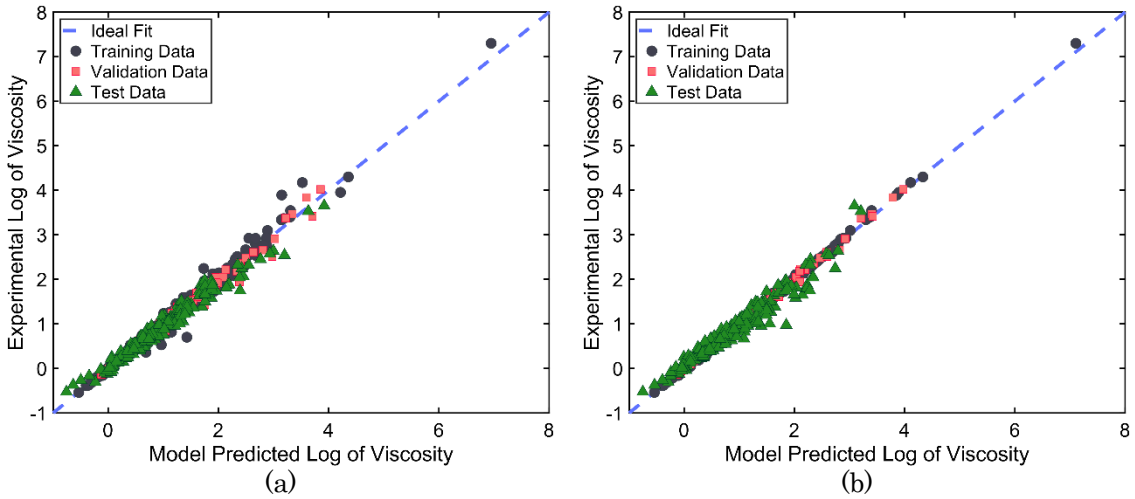


Figure 6.8: Model predicted viscosity obtained from the best models (Model III) with (a) static (b) dynamic descriptors for the training (black circles), validation (red squares), and test data (green triangles) sets. The blue dashed lines represent the ideal predictions.

Figures 5.8(a) and 5.8(b) show experimental viscosity and viscosity predicted using the best models with static and dynamic descriptors for all data sets over a wide range of temperatures. Note that the y-axis is $\log\eta(T)$. From the statistics in Table 6.3 and Figure 6.9, we can see that both models performed exceptionally well with only three or four predictors, including temperature. The model with three static descriptors and the model with three dynamic descriptors performed the same, with an average R^2 value of 97.7%.

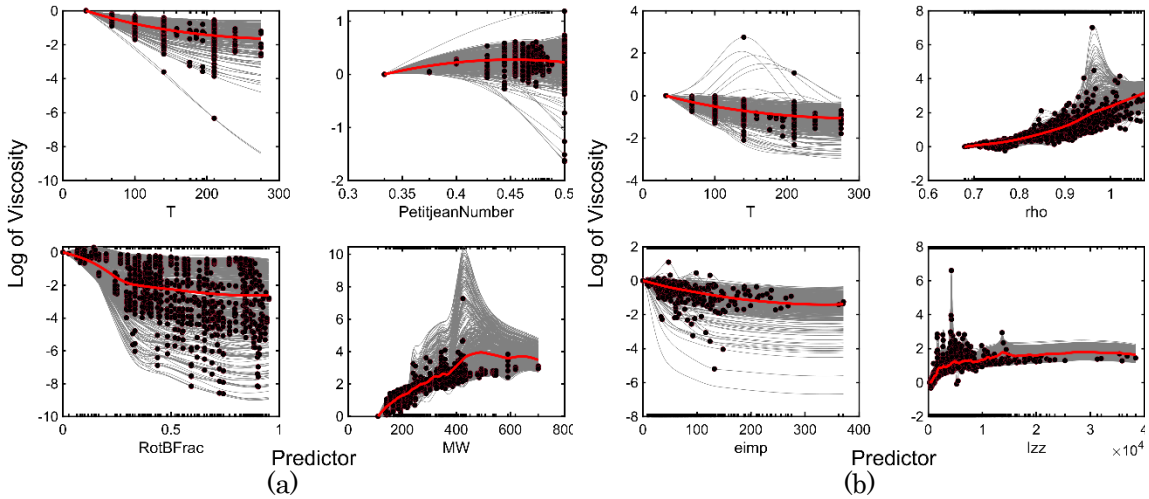


Figure 6.9: The partial dependency (red curves) and the individual conditional expectation (grey curves) of each predictor in the best viscosity models (Model III) with (a) static (b) dynamic descriptors. The scattered circular symbols represent the relationship between a predictor variable and viscosity for each observation. The plots are offset so that the viscosity starts from zero better to illustrate the cumulative effect of each predictor.

Like density, Figure 6.9 shows ICE plots (gray lines) and a PDP plot (red line) for each predictor. Figures 5.9(a) and 5.9(b) show the PDP and ICE plots of predictors in the best models with static (Equation (6.12)) and dynamic (Equation (6.15)) descriptors. The circle symbols are the predicted response by the predictor for each data point. PDPs of static descriptors show viscosity decreasing with increasing T and $RotBFrac$, increasing with $PetitjeanNumber$ and MW. PDPs of dynamic descriptors show that viscosity decreases with increasing T and $eimp$, but increases with the rho and I_{zz}. Similar finding were reported in a recent QSPR study [251] where $PetitjeanNumber$, $RotBFrac$, and MW were found to be correlated with viscosity. Many previous experimental and simulation studies reported a power-law relationship between MW and viscosity [231,289], consistent with the trend in Figure 6.9(b). Figures 5.10(a) and 5.10(b) show the average LIME values over all observations for the viscosity models with static and dynamic descriptors. The direction of the bars in the LIME plots indicate the same trend of the predictor with viscosity as the PDPs and ICE plots in Figure 6.9, i.e., viscosity decreases with increasing $RotBFrac$, T, but decreases with increasing $PetitjeanNumber$ and MW. The consistency between all three interpretation tools validates the various methods used to interpret the developed GPR models for establishing predictors relationship and importance to the response variable (density/viscosity).

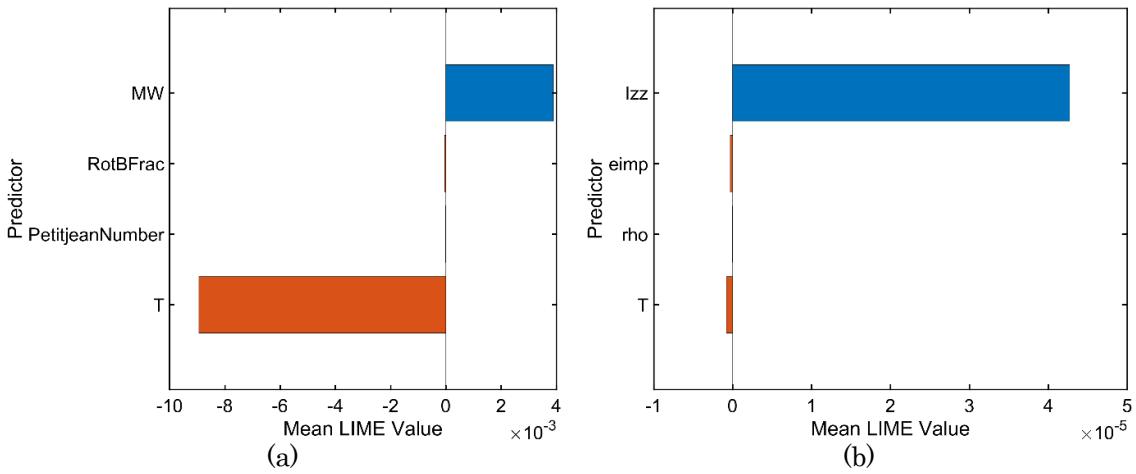


Figure 6.10: The average LIME value or coefficient for each predictor in best viscosity models (Model III) with the (a) static (b) dynamic descriptors. The orange and blue colors represent the negative and positive relationship between a predictor and the response variable. The size of a bar represents the overall importance of a predictor of a model.

Table 6.4 gives the R^2 values of viscosity models for the training, validation, and test data sets when a predictor was dropped from the best models with static and dynamic descriptors. The R^2 decreased by 0.6%, 12.3%, 29.1%, or 54.4% when $PetitjeanNumber$, $RotBFrac$, MW, or T term was dropped from the

best viscosity model with static descriptors. Therefore, the order of importance of the predictors in the model with static descriptors is $PetitjeanNumber < RotBFrac < MW < T$. Similarly, for the best viscosity model with dynamic descriptors, the order of predictor importance is $eimp < T < I_{zz} < rho$. When the $PetitjeanNumber$ is dropped from the model with static descriptors, the performance decreased by only 0.6% which suggested that this term could be removed permanently from the model. Therefore, the most important predictors of viscosity are T, MW, $RotBFrac$, $eimp$, I_{zz}, and rho. This finding contradicts a previous study where the $PetitjeanNumber$ was found to be highly correlated with viscosity [251,290], but may be explained by the fact that our models were developed for only pure hydrocarbons.

Table 6.4: Performance of the best viscosity models when a predictor was removed from the models.

Static Descriptors (Equation (6.12))			**Dynamic Descriptors** (Equation (6.15))		
Term Dropped	Avg. R^2	Drop in Avg. R^2	Term Dropped	Avg. R^2	Drop in Avg. R^2
None	97.7%	0.0%	None	97.7%	0.0%
$PetitjeanNumber$	97.1%	0.6%	$eimp$	90.2%	7.5%
$RotBFrac$	85.4%	12.3%	T	77.7%	20.0%
MW	68.6%	29.1%	I_{zz}	60.3%	37.4%
T	43.4%	54.4%	rho	53.7%	44.0%

We developed viscosity models with combined both static and dynamic descriptors. The combined models performed only slightly better (R^2 of 0.982) than the models with only static (R^2 of 0.977) or only dynamic (R^2 of 0.977) descriptors. The details of these models can be found in the supplementary information. The MATLAB code for the best density and viscosity models with static and dynamic descriptors are also provided in the supplementary information.

6.4. Conclusion

A GPR-based model was trained with Bayesian optimization to accurately predict the dynamic viscosity and density of complex hydrocarbons over a wide range of temperatures. We presented a top-down systematic approach to developing simple models using various robust ML algorithms. Our approach (1) removed redundant and strongly correlated predictor, (2) assessed the risk of overfitting and underfitting in models, (3) ensured that important predictors were included in the model, (4) assessed the quality of the model predictions, and (5) included model-agnostic interpretation. The best subset regression approach evaluated all combinations of significant predictors. Notably, the

developed models involved very few (less than or equal to five) and relatively simple predictors. The models were shown to be highly accurate in predicting experimental dynamic viscosity and density as a function of temperature for a variety of hydrocarbons. The R^2 values for the best density model were 99.1%, 98.9%, and 95.2% for the training, validation, and test data sets. The R^2 values for the best viscosity model were 99.1%, 98.9%, and 95.2% for the training, validation, and test data sets.

In addition, we investigated and compared the feasibility and versatility of using dynamic and static molecular descriptors to predict density and viscosity of hydrocarbons. Models with dynamic descriptors performed as well as or better than models with static descriptors, even though the pool of dynamic descriptors (157) was significantly smaller than that of the static descriptors (1444). For density, the dynamic descriptor-based models slightly outperformed static descriptor-based models, i.e., average R^2 values for the two models were 99.6% vs. 99.3%, respectively. For viscosity, models with static and dynamic descriptors were same with average R^2 values of 97.7%. This outcome suggests that descriptors obtained from MD simulations are more robust in capturing the dynamics of the molecules at various operating conditions. This is likely to be more important for other operating conditions, such as high pressures or shear rates, that may significantly affect molecular features. However, a further comparative study is needed with a larger pool of dynamic descriptors.

Model-agnostic interpretation was conducted by partial dependency plots, individual conditional expectation plots, LIME values, and by trimming predictor from the models. The model interpretation showed that the $RotBtFrac$, $piPC3$, rho, and R_g of the molecules are important predictors of density. Density is negatively correlated with $RotBtFrac$ and $evdwl$, but positively correlated with $piPC3$, rho, and R_g, For viscosity, model interpretation showed that $RotBtFrac$, MW, I_{zz}, rho, and $eimp$ are important predictors. Viscosity is positively related to MW, R_g, I_{zz} and rho, but negatively related to $RotBFrac$ and $eimp$. The consistency between the various interpretation methods confirms the validity of each method in evaluating predictors importance. Most importantly, the developed ML-based predictive models can be used to quickly predict the properties of existing hydrocarbons at other temperatures or for different hydrocarbons, which may enable the design of novel and innovative hydrocarbon molecules with tunable properties.

6.5. Supplementary Information

The following supplementary materials are provided on GitHub at https://github.com/panwarp/Supplementary-Materials which include:

- Schematics of all molecules.
- Definition of the molecular descriptors.
- All experimental data with the static and dynamic descriptors of all molecules.
- LAMMPS data files of the molecules with forcefield parameters and initially built atomic positions. LAMMPS input files to run the MD simulations.
- MATLAB files of the best models to predict temperature-dependent density and viscosity.
- Density and viscosity models with combined static and dynamic descriptors.
- Density models with dynamic descriptors excluding simulation-calculated density.

Chapter 7. Summary and Future Work

7.1. Summary

The mechanical efficiency of machines is affected by the viscosity of the fluid. Viscosity modifiers that thicken the fluid, therefore, play an important role in efficiency. Viscosity modifiers are believed to improve the mechanical efficiency of machines, such as hydraulic systems, partially by enabling formulation with lower molecular weight base oils. In Chapter 2, this concept was directly tested in a pump dynamometer using mixtures of low traction synthetic polyalphaolefin base oils, bis(2-ethylhexyl) adipate ester, and polyisobutylene. Lower viscosity fluids directly correlated to better mechanical efficiency but decreasing the viscosity of the synthetic base oil by adding viscosity modifier did not have the same effect. However, MD simulations showed that solution viscosity was directly correlated to elongation of the polymer under shear, which, together with calculations of the key shear rate range in a pump, suggested ways of designing viscosity modifiers to achieve a specific viscosity profile that maximizes mechanical efficiency.

Many application-relevant fluids exhibit shear thinning, where the viscosity decreases with a shear rate above some critical shear rate. For hydraulic fluids formulated with polymeric additives, the critical shear rate is a function of the molecular weight and concentration of the polymers. Chapter 3 presented a model for predicting the critical shear rate and Newtonian viscosity of fluids, intending to identify a fluid that shear thins in a specific range relevant to hydraulic pumps. The model was applied to predict the properties of fluids comprising polyisobutylene polymer and polyalphaolefin base oil. The theoretical predictions were validated by comparison to viscosities obtained from experimental measurements and MD simulations across many decades of shear rates. Results demonstrated that the molecular weight of the polymer plays a key role in determining the critical shear rate, whereas the concentration of polymer primarily affects the Newtonian viscosity. The simulations were further used to show the molecular origins of shear thinning and critical shear rate. The atomistic simulations and simple model developed in this work can ultimately be used to formulate polymer-enhanced fluids with ideal shear thinning profiles that maximize the efficiency of hydraulic systems.

The chemistry and structure of base oil and polymer additive molecules in lubricants directly affect key performance metrics such as viscosity index, thickening efficiency, and traction coefficient. However, the relationship between molecular properties and these metrics is still not fully understood, inhibiting the design of new fluids with potentially improved performance. In

Chapter 4, MD simulations were used to identify structure-property-function relationships for model lubricants consisting of branched and linear polymers having chemistries consistent with commercially available products. Four fluids were formulated with different polymers but with the same kinematic viscosity at 100 °C. Then, the simulation-calculated Newtonian viscosities at 40 and 100 °C, viscosity index, thickening efficiency, and traction coefficient at 40 °C were validated by direct comparison to experimental data. Next, the molecular origins of differences in the viscosity index, thickening efficiency, and traction coefficient between the fluids were investigated by calculating the multiple structural properties of the polymers in the simulations. Finally, the simulations were used to develop simple empirical models using best subset linear regression analysis to rapidly predict viscosity index, thickening efficiency, and traction coefficient. The atomistic simulations and empirical models developed in this work can ultimately guide the design of new lubricants or additives.

Although the development of novel materials has benefited from MD simulations, the efficiency and cost of these methods still limit the size and time scale of the materials and processes that can be studied. In addition, experimental measurements are not always possible for molecules generated *in silico* because they must be synthesized first which can be expensive and time-consuming. In such cases, quantitative-structure-property-relationship (QSPR) models developed using advanced ML algorithms are robust and allow rapid and reliable prediction of material properties and facilitates innovative materials design. However, a suitable set of molecular descriptors must be used to develop a simple and reliable QSPR model. For this purpose, there was an urgent need for a comprehensive and flexible toolkit to integrate with MD simulations. Therefore, in Chapter 5, we developed an open-source Python-based post-processing tool called PyL3dMD. PyL3dMD is compatible with the popular MD simulation package LAMMPS and enables users to compute nearly 3D (dynamic) molecular descriptors. The availability of this open-source post-processing tools to analyze simulation trajectories and relate 3D conformation to properties is key to helping researchers conduct advanced QSPR and inverse QSPR studies.

This newly developed python package was implemented and validated in Chapter 6 by calculating dynamic descriptors of 305 complex hydrocarbons from their LAMMPS MD simulation trajectory and data files. These dynamic descriptors were then used to develop ML-based QSPR models for predicting temperature-dependent viscosity and density of hydrocarbons. Lubricants are formulated by blending base oils with additives to enhance properties and meet performance specifications. A lubricant can contain up to 98% base oil which are mostly complex hydrocarbons [8,232]. The molecular composition of a base

oil directly affects its properties and, therefore, the performance of a lubricant [233]. Therefore, in Chapter 6, we developed Gaussian process regression-based models to accurately predict the temperature-dependent density and dynamic viscosity of 305 complex hydrocarbons. In the process, strongly correlated/collinear predictors were trimmed, important predictors were selected by least absolute shrinkage and selection operator (LASSO) regularization and prior domain knowledge, hyperparameters were systematically optimized by Bayesian optimization, and the models were interpreted. The approach provided versatile and quantitative structure–property relationship (QSPR) models with relatively simple predictors for determining the dynamic viscosity and density of complex hydrocarbons at any temperature. In addition, we developed MD simulation-based descriptors and evaluated the feasibility and versatility of dynamic descriptors from simulations for predicting material properties. It was found that the models developed using a comparably small pool of dynamic descriptors performed similarly in predicting density and viscosity than models based on many more static descriptors. The best models were shown to predict density and dynamic viscosity with coefficient of determination (R^2) values of 99.6% and 97.7%, respectively, for all data sets, including a test data set of 45 molecules. Finally, partial dependency plots (PDPs), individual conditional expectation (ICE) plots, local interpretable model-agnostic explanation (LIME) values, and trimmed model R^2 values identified the most important static and dynamic predictors of density and viscosity.

7.2. Future Work

7.2.1. Improve Developed Python Package - PyL3dMD

Currently version of our developed python package, PyL3dMD, has some limitations and calculates nearly 2000 descriptors using LAMMPS data and trajectory files. Therefore, in future, we would like to add the following features and functionality:

- Include more complex descriptors such as persistence length, orientation tensor, and solvent accessible surface area.
- The current package is only compatible with the LAMMPS MD simulation tool. Therefore, we would like to expand it so that it can be used to calculate dynamic descriptors using trajectories of molecules generated by other commonly used MD simulation tools.
- The current package assumes that the LAMMPS simulation is in "real" units (mass = grams/mole, length = Angstroms, time = femtoseconds, energy = Kcal/mole, temperature = Kelvin, pressure = atmospheres, electric field = volts/Angstrom, density = gram/cm^3); therefore, in our next version we will ensure that it can be used for any unit system.
- The current version uses the rdkit [214] python package to calculate adjacency and distance matrices using the *.mol file of a molecule. In the next version, we will ensure that the package can perform all the task standalone.
- The current package uses multiprocessing and linear algebra to a great extent. We will improve computational efficiency by making changes in the script such as replacing loops with linear/matrix algebra, incorporating list comprehensions, using join() instead of plus symbol to concatenate strings, reducing the number of times functions are called, and by exiting loops and functions as soon as the desired outcome is achieved.
- Include machine learning tools to provide the user the option of quickly developing simple predictive QSPR models using our ML approach described in Chapter 6.

7.2.2. Develop ML-Based QSPR for Predicting Viscosity, VI, TE, and TC Model of any Polymer-enhanced Fluid

In our third study, we simulated viscosity and traction coefficient and then used simulation data to develop simple empirical models to predict VI, TE, and TC. This study only included four fluids of similar/fixed kinematic viscosity (FKV) that had different polymers but the same base oil. As shown in the study, the models predicted the properties of the original four fluids used in

the model development process with good accuracy. However, the models could not accurately predict all properties of the fifth fluid (F-ROMP) that had different viscosity base oil and polymer. This suggested the models should be further improved to be used for the fluids formulated with a polymer of any degree of branching, any concentration of polymer, or/and any viscosity base oil.

To improve the performance of our models in predicting properties of any formulation of fluid accurately, we will formulate fluids of varying viscosity base oils with polymers of different chemistries and structures. In this study, we plan to include wide range of fluid formulation, including fluids formulated using same concentration of polymers and to have similar/same viscosity index. Then, we will measure the Newtonian viscosity of the fluids at temperatures of 20, 30, 40, 50, 60, 70, 80, 90, and 100 °C using a viscometer and rheometer. Also, the TC of fluids at temperatures of 40, 60, 80, and 100 °C and slide-to-roll ratios of 2 to 60% will be measured using a mini-traction-machine. Then, using the MD simulations and our python package, PyL3dMD, we will calculate dynamic descriptors of the fluids at the operating conditions consistent to experimental measurements. These dynamic descriptors then predict measured properties (viscosity, VI, TE, and TC) of the fluids using the ML approach presented in Chapter 6. The new models will again be evaluated using a set of test fluids that were not used in the model development process to check the ability of models to predict VI, TE, and TC of any formulation. Ultimately, these ML-based QSPR models will be used to accurately predict properties of molecules not yet synthesized.

7.2.3. Develop a Multiple Inputs and Multiple Outputs Neural Network for Predicting Multiple Properties

We would like to make use of available large experimental data set of important bulk properties in research articles and scientific data banks, as we did in Chapter 6, to develop a multiple-input and multiple-output neural network for predicting all these properties of lubricants from a single model, as shown in Figure 7.1. In this model, molecular descriptors of lubricants and operating conditions at which experimental data were conducted will be provided as the inputs to the neural network.

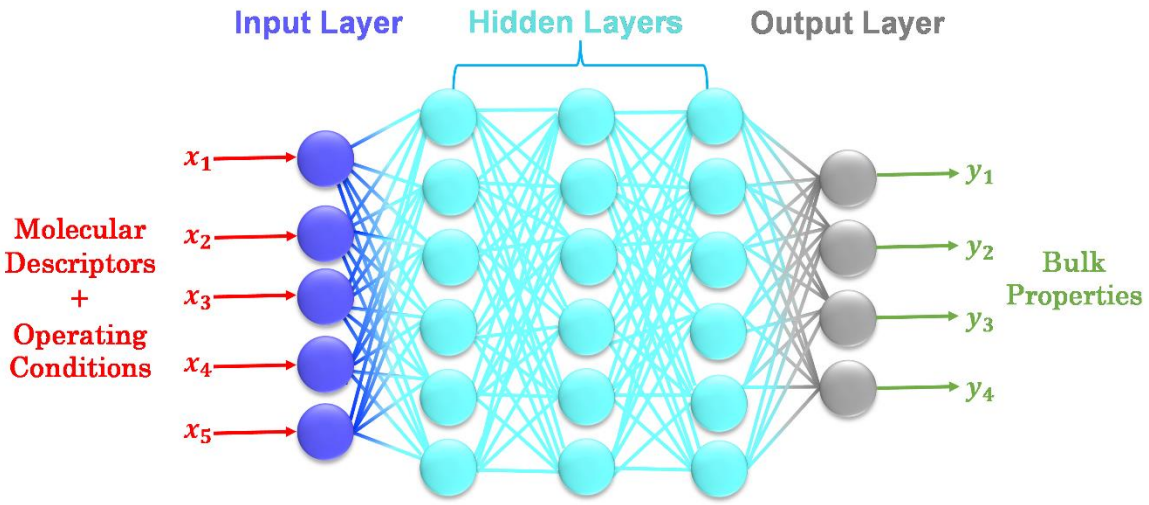


Figure 7.1: Multiple-input and multiple-output neural network.

To achieve optimum machine performance, it is necessary to synergically understand the effects of composition and chemistries of lubricants on their key performance metrices and properties because most of the properties are related to each other in one way or other. For example, the Sieder–Tate equation [44], Equation (7.1), relates the coefficient of heat transfer (h) to thermal conductivity (K), density (ρ), specific heat (C_p) at constant pressure, and Newtonian viscosity (η_0).

$$h \propto \frac{K^{0.67}\rho^{0.8}C_p^{0.33}}{\eta_0^{0.47}} \tag{7.1}$$

The cooling performance of lubricant plays an important role in the energy efficiency of a system. The heat generated from the friction causes the temperature of the oil film to increase. This higher temperature reduces the viscosity of the oil and thickness of the lubricating film, which ultimately results in inefficiency and wear related failures [4]. Fluids with higher specific heat can contain large amounts of heat. Fluids with higher thermal conductivity are better at carrying heat away from the tribological contacts to dissipate effectively [44]. Also, high thermal conductivity and high specific heat are good for film formation, especially under high-slip and high-load conditions. High heat capacity implies that the fluid increases its vibrational and rotational quantum state instead of increasing its bulk temperature [291,292]. The high thermal conductivity increases the amount of energy

transferred from the contact out to the surrounding material and fluid, thereby lowering the temperature in the contact zone and increasing viscosity [291].

The molecular structure in fluids determines the thermal capacity and thermal conductivity of a lubricant [291]. For instance, the thermal stability of the base oil improves from Group I to Group IV, that is, with increasing group numbers [44,293]. The molecular structure determines how many "quantum states" (that is, how many ways it can freely rotate or vibrate) it can have. The higher the number of the rotational and vibrational quantum states, the higher the thermal capacity [294]. It takes higher energy input to raise the averaged kinetic energy, therefore, the temperature when there are high numbers of vibrational and rotational states. The thermal conductivity of base oil was correlated to the fluid's molecular diffusivity, which is highly dependent on molecular structure and chemistry [292]. It has been observed that additives in base oils can modify cooling performance [295]. The longer chains of molecules providing better cooling performance [295].

Therefore, we would like to develop a single model for predicting properties such as density, Newtonian viscosity, shear viscosity (shear thinning/shear stability), viscosity index, thickening efficiency, traction coefficient, specific heat, and thermal conductivity for a wide range of fluid formulations and over a wide range of operating conditions.

7.2.4. Inverse QSPR - Formulate and Evaluate Energy Efficient Hydraulic Fluids

The objectives of achieving reduced emissions, higher fuel economy, and longer service (drain) intervals create a need for the industry to develop advanced lubricants, the quality of which is defined by the new lubricant performance specifications (For instance, a hydraulic fluid should have a viscosity range of 25-35 cSt at 40 °C, viscosity range at 5-10 cSt at 100 °C, high viscosity index (>170), high thermal conductivity, high specific heat, optimum shear thinning, low polymers treat rate (high thickener polymeric additive), and low traction coefficient). Apart from the obvious use of QSPR models to predict the properties of both existing and new molecules, the use of QSPR models to virtually synthesize new molecules is currently a high-interest subject [296]. Therefore, we would like to extend our work by implementing the outcomes of modeling key performance matrices of lubricants and additives to formulating lubricants and additives with optimum performance matrices and demonstrate the uses of the developed ML-based QSPR models and MD simulations. This process of reverse engineering structure of molecules to achieve a desired property is known as inverse QSPR, as shown in Figure 7.2.

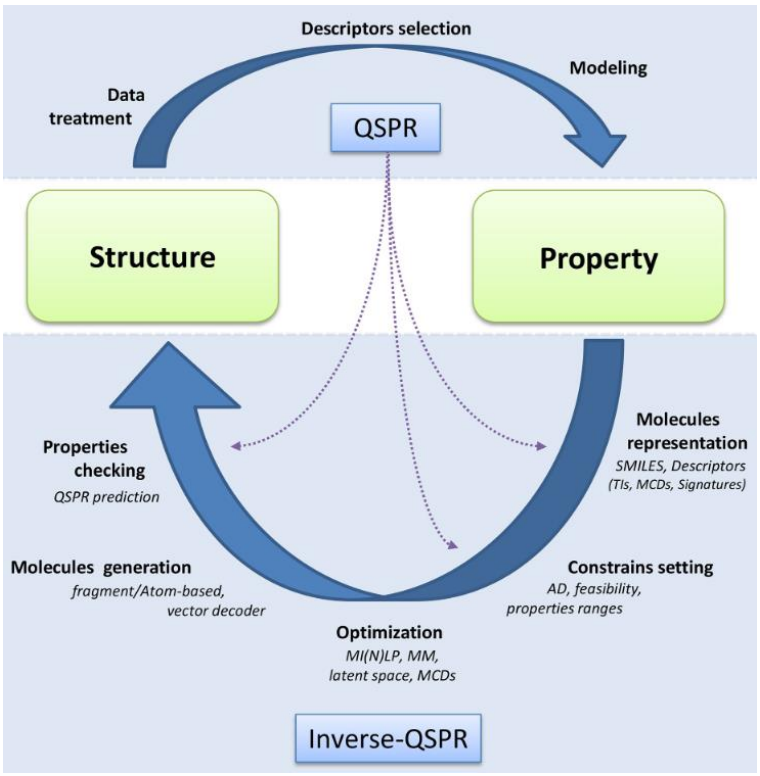


Figure 7.2: Inverse quantitative-structure-property-relationship [296].

Knowing that each application has specific requirements for lubricant properties with enhanced performance, we will focus specifically on hydraulic pumps for this research stage, even though the simple predictive models can be used for any application. The total power loss in an axial piston pump can be primarily divided into (1) volumetric losses and (2) mechanical losses. The volumetric losses account for the gaps flow losses - leakage flow occurs from the micron-level lubricating gaps due to the pressure difference across these gaps, cross-port flow losses - back flow from the outlet port to the inlet port forms the internal leakage, which is mainly due to a design flaw, and compressibility flow losses - loss of volume of fluid due to its compressibility. The mechanical losses account for the torque required to overcome the viscous friction in lubricating interfaces, mechanical friction in lubricating interfaces, friction in roller bearings and sealing components, churning losses - the losses caused by the rotation of the pump assembly in the case filed by fluid, and timing of pressure and load sensing valves [30,39,297]. Figure 7.3 shows the distribution of major power losses in pumps at two temperatures, 50 and 70°C, two swash plate angles of 10 and 100%, two speeds of 700 and 2000 rpm, and two system pressure of 50 and 170 bar.

Knowing the losses and their causes (leakage or friction), as shown in Figure 7.3, we can formulate fluids with optimized viscosity at different operating conditions to obtain maximum efficiency. This would be done by varying the parameters that affect bulk properties of the lubricants - back bone length, branching size/number, size of molecules, polarity, saturation/unsaturation, and ring. Generally, a hydraulic fluid should have a viscosity range of 25-35 cSt at 40 °C, viscosity range at 5-10 cSt at100 °C, high viscosity index (>170), high thermal conductivity, high specific heat, optimum shear thinning, low polymers treat rate (high thickener polymeric additive), and low traction coefficient.

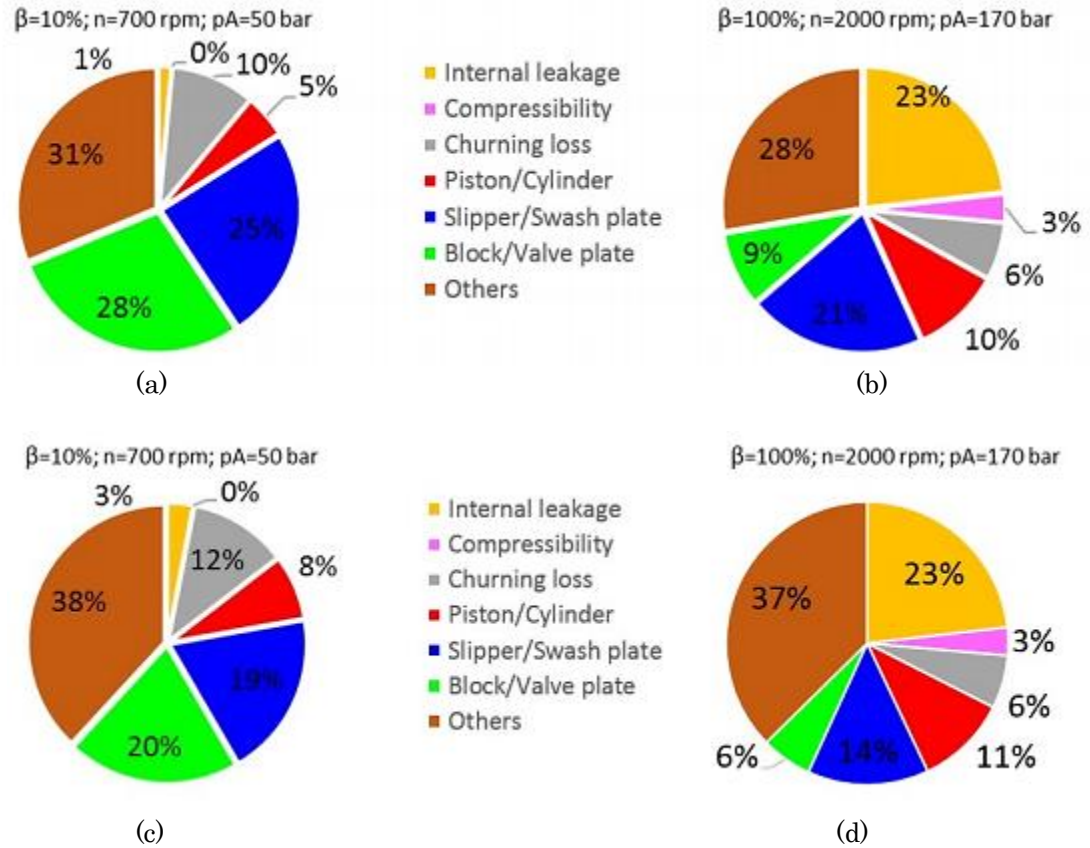


Figure 7.3 Distribution of power losses at (a) 50 °C and low power operating condition, (b) 50 °C and high power operating condition, (c) 70 °C and low power operating condition, (d) 70 °C and high power operating condition [297].

After formulating a set of fluids (2-3 fluids), we would evaluate our fluids by measuring important fluid properties such as viscosity using a rheometer, specific heat using a differential scanning calorimeter, thermal conductivity using a hot wire method, and traction coefficients using a tribometer at various operating conditions. Finally, we would like to evaluate an axial piston pump

performance in a dynamometer with new fluids of potentially improved performance.

7.3. Concluding Remarks

In conclusion, molecular dynamics simulations, machine learning, and quantitative-structure-property-relationship modeling were used to investigate the influence of fluid molecules on their key lubricating properties. These key properties include density, Newtonian viscosity, shear viscosity, viscosity index, thickening efficiency, and traction coefficient at wide ranges of operating conditions. In this research work, we developed new MD simulation techniques to accurately simulate properties of commercial grades polymer-enhanced fluids, various empirical and simple ML-based models to understand the difference in the properties of the fluids regarding chemistry and structure of their molecules. We presented a model to formulate polymer-enhanced fluids with ideal shear thinning profiles that could maximize the efficiency of hydraulic systems such as pumps. We also developed a ML approach to develop simple but highly accurate models and conducted a detailed model-agnostic interpretation to study the effect of molecular features of lubricants individually on their properties. In addition, we also developed a one-of-a-kind python package that could be used to determine hundreds of different molecular features/descriptors by post-processing MD simulation output files. These molecular descriptors then can be used to establish quantitative-structure-property-relationship for any application and material.

Future work includes investigating the synergistic effect of polymeric additives and base oils on various bulk properties of lubricants by developing single multiple-input and multiple-output neural network using the techniques and tools developed in this research work. This neural network then can be used to reverse engineer new lubricants or additives with desired properties. Ultimately, we plan to demonstrate the effectiveness of the developed tools, techniques, and models by formulating hydraulic fluids with predefined fluid properties for optimum efficiency across a range of operating conditions.

We believe that advancements in all these areas can lead to the design of advanced lubricants and lubricated systems with improved lubrication performance, component reliability, and energy efficiency which, in turn, may lead to technologies that reduce the global energy consumption that originates from tribological contacts.

Abbreviations and Nomenclatures

ABBREVIATIONS

ANN	artificial neural network
ASTM	American society for testing and materials
ANOVA	analysis of variance
API	American petroleum institute
BDIP	butadiene isoprene
CHARMM	chemistry at Harvard macromolecular mechanics
CODESSA	comprehensive descriptors for structural and statistical analysis
COSMO-RS	conductor like screening model for real solvents
DEHA	bis(2-ethylhexyl) adipate ester
EHL	elastohydrodynamic lubrication
EJ	exajoules
EMD	equilibrium molecular dynamics
EoS	equations of states
FKV	fixed kinematic viscosity
F-ROMP	user defined fluid name
FTR	fixed treat rate
FVI	fixed viscosity index
GC	group-contributions
GK	Green–Kubo
GNP	gross annual product
GPR	gaussian process regression
HL	hydrodynamic lubrication
HSD	hydrogenated styrene-diene
HV46/15	user defined fluid name
ICE	individual conditional expectation
LAMMPS	large-scale atomic/molecular massively parallel simulator
LASSO	least absolute shrinkage and selection operator
LIME	local interpretable model-agnostic explanation
MABD	methacrylate butadiene
MD	molecular dynamics
ML	machine learning
MLR	multiple linear regression
MSE	mean squared error
MTM	mini-traction-machine
NEMD	non-equilibrium molecular dynamics
NPT	isothermal–isobaric ensemble
NVE	canonical ensemble
NVT	microcanonical ensemble
OCP	olefin copolymer
OPLS	optimized potentials for liquid simulations
PAMA	polyalkyl methacrylate
PAO	polyalphaolefin
PAO2	polyalphaolefin of 2 mm^2/s at 100 °C
PAO4	polyalphaolefin of 4 mm^2/s at 100 °C
PAO8	polyalphaolefin of 8 mm^2/s at 100 °C
PCA	principal component analysis

PDI	polydispersity index
PDP	partial dependency plot
PIB	polyisobutylene
PIB1300	polyisobutylene of 1300 g/mol
PIB6000	polyisobutylene of 6000 g/mol
PLS	partial least-squares regression
PR	polynomial regression
PyL3dMD	python LAMMPS 3D molecular descriptors package
QSPR	quantitative structure-property relationships
RMSE	root mean squared error
ROMP	user defined polymer name
SBCP	styrene block copolymers
SEM	standard errors of the mean
SMILES	simplified molecular input line entry system
TC	traction coefficient
TE	thickening efficiency
TraPPE	transferable potentials for phase equilibria
UA	united atom
VG	viscosity grade
VI	viscosity index
VIF	variance inflation factor
VII	viscosity index improver
VM	viscosity modifier

NOMENCLATURES

$AATS2e$	Broto-Moreau autocorrelation-lag 2/weighted by Sanderson electronegativity
a	Yasuda parameter
b	asphericity
C	concentration of polymer
$\mathbf{C}$	correlation matrix or matrix of the correlation coefficient
C_p	specific heat at constant pressure
c	acylindricity
c_p	concentration of polymer
c_s	concentration of solvent or base oil
$\partial u_x/\partial z$	flow gradient or momentum flux
$eimp$	energy due to improper interaction
$evdwl$	energy due to van der Waals interactions
$\dot{\gamma}$	shear rate
$\dot{\gamma}_{cr}$	critical shear rate
h_{min}	minimum film thickness
I_{zz}	diagonal component of moment of inertia tensor
i	i^{th} data point
K	bulk modulus
K	thermal conductivity
$Kier1$	first kappa shape index
k	relative shape anisotropy
k_B	Boltzmann constant
ke	kinetic energy

L_x, L_y, L_z	molecules length in x, y, and z direction
$\log \eta(T)$	logarithmic of temperature-dependent dynamic viscosity
λ	time constant or rotational relaxation time
λ_x, λ_y, λ_z	eigen values/principal moments of gyration tensor
M	molecular weight
M_n	number average molar mass of polymer
M_w	mass average molar mass of polymer
M_{wb}	molecular weight of backbone
M_{wp}	percentage molecular weight of polymer in backbone
M_{wt}	total molar mass or molecular weight of polymer
MW	molecular weight of molecule
μ	dynamic/absolute viscosity
μ_0	viscosity at zero shear
μ_∞	viscosity at infinite shear
u_x	velocity of atoms in the x-direction
N	total number of data points
n	power-law exponent
η	dynamic viscosity
η	shear viscosity of fluid
$\eta(T)$	temperature-dependent dynamic viscosity
η_0	Newtonian viscosity of fluid
η_{0B}	Newtonian viscosity of base oil
η_{0s}	Newtonian viscosity of fluid
η_∞	viscosity at infinite shear rate of fluid
υ	kinematic viscosity
v	velocity
ν_{0s}	Newtonian kinematic viscosity of the polymer-containing fluid
ν_{0B}	Newtonian kinematic viscosity of base oil
η_{0p}	Newtonian dynamic viscosity of polymer
η_{0s}	Newtonian dynamic viscosity of solvent or base oil
ω	rotational frequency
$P_{\alpha ß}$	pressure or stress tensor
$PetitjeanNumber$	Petitjean number
P_{xz}	shear stress
p	pressure
$piPC3$	conventional bond order ID number of order 3
Q	flow rate
Q_i	theoretical flow rate
Q_k	compressible flow losses
Q_l	pressure driven flow losses
R	end-to-end distance
R^2	coefficient of determination or R-squared
$R_{a,b}$, $R_{a,J}$	average or RMS roughness of the two surfaces in contact
R_j^2	R^2 of predictor j on the remaining predictors
$RotBFrac$	fraction of rotatable bonds, excluding terminal bonds
R_g	radius of gyration
R_g	universal gas constant
$R_{g,xx}^2$, $R_{g,yy}^2$, $R_{g,zz}^2$	diagonal elements of gyration tensor
$R_{g,xy}^2$, $R_{g,xz}^2$, $R_{g,yz}^2$	off-diagonal elements of gyration tensor

rho	MD simulation-calculated density from NPT ensemble
ρ	density or mass density
ρ_B	density of solvent or base oil
ρ_p	density of polymer
ρ_s	density of solvent or base oil
$\rho(T)$	temperature-dependent density
σ	standard deviation
T	temperature
T	temperature in °F
T	torque
T_e	effective or measured torque
T_i	theoretical torque
T_l	viscous torque losses
T_t	turbulent torque losses
t	time
τ	shear stress
θ	temperature
V	volume
V_i	displacement
Wi	Weissenberg number
$\bar{y}$	mean experimental value of the response variable
y_i^{exp}	experimental value of the response variable
y_i^{pred}	model predicted value of the response variable

Bibliography

[1] HP, J., 1966, "Lubrication (Tribology)–A Report on the Present Position and Industry's Needs.," London HM Station. Off.

[2] I.M. Hutchings, 2016, "Leonardo Da Vinci's Studies of Friction," Wear, **360–361**, pp. 51–66.

[3] Peter Jost, H., 1990, "Tribology-Origin and Future," Wear, **136**(136), pp. 1–17.

[4] Holmberg, K., and Erdemir, A., 2019, "The Impact of Tribology on Energy Use and CO2 Emission Globally and in Combustion Engine and Electric Cars," Tribol. Int., **135**, pp. 389–396.

[5] Holmberg, K., and Erdemir, A., 2017, "Influence of Tribology on Global Energy Consumption, Costs and Emissions," Friction, **5**(3), pp. 263–284.

[6] Lee, P. M., and Carpick, R., 2017, *Tribological Opportunities for Enhancing America's Energy Efficiency*, A report to the Advanced Research Projects Agency-Energy (ARPA-E) at the US Department of Energy 14.

[7] Woydt, M., Gradt, T., Hosenfreldt, T., Luther, R., Rienäcker, A., Wetzel, F.-J., and Wincierz, C., 2019, *Tribology in Germany: Interdisciplinary Technology for the Reduction of CO2-Emissions and the Conservation of Resources*.

[8] Rizvi, S. Q. A., 2009, *A Comprehensive Review of Lubricant Chemistry, Technology, Selection, and Design*, ASTM International, Conshohocken.

[9] Hilton, M. R., and Fleischauer, P. D., 1992, "Applications of Solid Lubricant Films in Spacecraft," Surf. Coatings Technol., **54–55**, pp. 435–441.

[10] Stachowiak, G., and Batchelor, A., 2006, *Engineering Tribology*, Burlington.

[11] Bair, S., 2019, *High Pressure Rheology for Quantitative Elastohydrodynamics*, Elsevier.

[12] Michael, P. W., Mettakadapa, S., and Shahahmadi, S., 2016, "An Adsorption Model for Hydraulic Motor Lubrication," J. Tribol., **138**(1).

[13] Ganesan, V., 2012, *Internal Combustion Engines*, McGraw Hill Education (India) Pvt Ltd.

[14] Bushan, B., 2001, *Modern Tribology Handbook*, CRC Press, Boca Raton.

[15] Wedeven, L. D., 1975, "What Is Ehd?," J. Am. Soc. Lubr. Eng., **31**, pp. 291–296.

[16] Dawson, D., and Higginson, G. R., 1966, *Elasto-Hydrodynamic Lubrication: The Fundamentals of Roller and Gear Lubrication*, Pergamon Press, England.

[17] Totten, G. E., Westbrook, S. R., and Shah, R. J., 2003, "Fuels and Lubricants Handbook: Technology, Properties, Performance, and Testing," Am. Soc. Test. Mater.

[18] Masjedi, M., and Khonsari, M. M., 2014, "Theoretical and Experimental Investigation of Traction Coefficient in Line-Contact EHL of Rough Surfaces," Tribol. Int., **70**, pp. 179–189.

[19] Khonsari, M. M., and Booser, E. R., 2008, *Applied Tribology: Bearing Design and Lubrication: Second Edition*, John Wiley & Sons.

[20] Fitch, E. C., 1992, *Proactive Maintenance for Mechanical Systems*, Elsevier Science, England.

[21] Mary, C., Philippon, D., Lafarge, L., Laurent, D., Rondelez, F., Bair, S., and Vergne, P., 2013, "New Insight into the Relationship between Molecular Effects and the Rheological Behavior of Polymer-Thickened Lubricants under High Pressure," Tribol. Lett., **52**(3), pp. 357–369.

[22] Larsson, R., Kassfeldt, E., Byheden, Å., and Norrby, T., 2001, "Base Fluid Parameters for Elastohydrodynamic Lubrication and Friction Calculations and Their Influence on Lubrication Capability," J. Synth. Lubr., **18**(3), pp. 183–198.

[23] Pensado, A. S., Comuñas, M. J. P., and Fernández, J., 2008, "The Pressure-Viscosity Coefficient of Several Ionic Liquids," Tribol. Lett., **31**(2), pp. 107–118.

[24] Ramasamy, U. S., Bair, S., and Martini, A., 2015, "Predicting Pressure-Viscosity Behavior from Ambient Viscosity and Compressibility: Challenges and Opportunities," Tribol. Lett., **57**(2).

[25] Bair, S., and Michael, P., 2010, "Modeling the Pressure and Temperature Dependence of Viscosity and Volume for Hydraulic Fluids," Int. J. Fluid Power, **2**, pp. 37–42.

[26] Bercea, M., Bercea, I., Nélias, D., and Flamand, L., 2002, "Polyethylene as an Additive for Mineral Oils- Part II: EHL Traction Behavior," Tribol. Trans., **45**(2), pp. 145–152.

[27] Wang, Y. S., Yang, B. Y., and Wang, L. Q., 2004, "Investigation into the Traction Coefficient in Elastohydrodynamic Lubrication," *TriboTest*, pp. 113–124.

[28] Sullivan, W. T., Oumar-Mahamat, H., Webster, M. N., and Brandes, E. B., 2010, "Lubricating Fluids with Low Traction Characteristics."

[29] Miller, M. K., Khalid, H., Michael, P. W., Guevremont, J. M., Garelick, K. J., Pollard, G. W., Whitworth, A. J., and Devlin, M. T., 2014, "An Investigation of Hydraulic Motor Efficiency and Tribological Surface Properties," Tribol. Trans., **57**(4), pp. 622–630.

[30] Michael, P. W., and Mettakadapa, S., 2016, "Bulk Modulus and Traction Effects in an Axial Piston Pump and a Radial Piston Motor," 10th Int. Fluid Power Conf., pp. 301–312.

[31] Sniderman, D., 2017, "The Chemistry and Function of Lubricant Additives," Tribol. Lubr. Technol., **73**(11), pp. 18–28.

[32] Wong, V. W., and Tung, S. C., 2018, "Friction, Lubrication, and Wear of Internal Combustion Engine Parts," Frict. Lubr. Wear Technol., **18**, pp. 899–915.

[33] Dorinson, A., and Ludema, K. C., 1985, *Mechanics and Chemistry in Lubrication.*, Elsevier.

[34] McGuire, N., and Micheal, P., 2016, "Efficient Hydraulic Systems Deliver the Power," Tribol. Lubr. Technol., **72**(5), pp. 36–42.

[35] Michael, P. W., Khalid, H., and Wanke, T., 2012, "An Investigation of External Gear Pump Efficiency and Stribeck Values," SAE Tech. Pap., **8**.

[36] Mettakadapa, S., Bair, S., Aoki, S., Kobessho, M., Carter, L., Kamimura, H., and Michael, P. W., 2015, "A Fluid Property Model for Piston Pump Case Drain and Pressure Compensator Flow Losses," ASME/BATH 2015 Symp. Fluid Power Motion Control. FPMC 2015.

[37] Görlitzer, H., Alibert, M., Herzog, S. N., and Neveu, C. D., "Dependence of Pump Flow Rate on the Viscosity of High VI Hydraulic Oils," *Proceedings of the 15th International Colloquium Tribology, Stuttgart/Ostfildern, Germany, 17–19 January 2006*, Technische Akademie Esslingen: Ostfildern, Germany, 2006.

[38] Devlin, M. T., Senn, J., Turner, T. L., Milner, J., and Jao, T., 2006, "Reduction in Axle Oil Operating Temperatures by Fluids with Optimized Torque Transfer Efficiencies," pp. 7–23.

[39] Michael, P. W., Garcia, J. M., Bair, S. S., Devlin, M. T., Michael, P. W., Garcia, J. M., Bair, S. S., Devlin, M. T., Michael, P. W., Garcia, J. M., Bair, S. S., Devlin, M. T., Martini, A., and Chemical, A., 2012, "Lubricant Chemistry and Rheology Effects on Hydraulic Motor Starting Efficiency Lubricant Chemistry and Rheology Effects on Hydraulic Motor Starting Efficiency," **2004**.

[40] Michael, P., Cheekolu, M., Panwar, P., Devlin, M., Davidson, R., Johnson, D., and Martini, A., 2018, "Temporary and Permanent Viscosity Loss Correlated to Hydraulic System Performance," Tribol. Trans., **61**(5), pp. 901–910.

[41] Strate, G. Ver, 1991, "Polymers as Lubricating-Oil Viscosity Modifiers," pp. 256–272.

[42] Cai, G., Zhang, L., Ma, L., and Eli, W., 2015, "Synthesis and Characterization of Polybutylacrylate Viscosity Index Improver with Anti-Wear Function," Lubr. Sci., **27**(4), pp. 209–216.

[43] Martini, A., Ramasamy, U. S., and Len, M., 2018, "Review of Viscosity Modifier Lubricant Additives," Tribol. Lett., **66**(2).

[44] Rizvi, S. Q. A., 2020, *A Comprehensive Review of Lubricant Chemistry , Technology , Selection , A Comprehensive Review of Lubricant Chemistry , Technology , Selection , And*.

[45] Koehler, B., and Jackson, J. M., 2010, "Evaluation of Hydraulic Efficiency Using High-Shear Viscosity Fluids," *SAE Technical Paper Series*.

[46] Neveu, C. D., Alibert, M. J. and Camera, F., 2010, "Impact of Fresh and Sheared Oil Viscosity Requirements on the Formulation of Hydraulic Fluids," goriva i Maz., **49**(2), p. 161.

[47] GOI, T., TANAKA, H., NAKASHIMA, K., and WATANABE, K., 2010, "Study on Stability of High Speed Traction Drive CVT for Aircraft Generator," J. Japan Soc. Aeronaut. Sp. Sci., **58**(678), pp. 203–209.

[48] Bell, D D, Raynard, A. E., 1980, "Design Study of Toroidal Traction Cvt for Electric Vehicles."

[49] Balakin, P. D., and Stripling, L. O., 2017, "Power and Energy Ratios in Mechanical CVT Drive Control," J. Phys. Conf. Ser., **858**(1).

[50] LaFountain, A., Johnston, G. J., and Spikes, H. A., 1998, "Elastohydrodynamic Friction Behavior of Polyalphaolefin Blends," Tribol. Ser., **34**, pp. 465–475.

[51] Covitch, M. J., and Trickett, K. J., 2015, "How Polymers Behave as Viscosity Index Improvers in Lubricating Oils," Adv. Chem. Eng. Sci.

[52] Gold, P. W., Schmidt, A., Dicke, H., Loos, J., and Assmann, C., 2001, "Viscosity-Pressure-Temperature Behaviour of Mineral and Synthetic Oils," J. Synth. Lubr., **18**(1), pp. 51–79.

[53] Moughon, L., and Wong, V. W., 2005, "Effects of Lubricant and Piston Design on Reciprocating Engine Friction," *2005 Fall Technical Conference of the ASME Internal Combustion Engine Division*, pp. 653–662.

[54] ASTM D2270, 2016, "Standard Practice for Calculating Viscosity Index from Kinematic Viscosity at 40 °C and 100 °C," ASTM Int.

[55] Stöhr, T., Eisenberg, B., and Müller, M., 2009, "A New Generation of High Performance Viscosity Modifiers Based on Comb Polymers," SAE Int. J. Fuels Lubr., **1**(1), pp. 1511–1516.

[56] Glass, J. E., Schulz, D. N., and Zukoski, C. F., 1991, *Polymers as Rheology Modifiers*, American Chemical Society, Washington.

[57] Canter, N., 2011, "Viscosity Index Improvers," Tribol. Lubr. Technol., **67**(9), pp. 10–22.

[58] Ahmed, N. S., and Nassar, A. M., 2011, *Lubricating Oil Additives*, Open Access Publisher.

[59] Ramasamy, U. S., Cosimbescu, L., and Martini, A., 2015, "Temperature-Dependent Conformations of Model Viscosity Index Improvers," Tribol. Lubr. Technol., **71**(5), pp. 30–31.

[60] Ramasamy, U. S., Lichter, S., and Martini, A., 2016, "Effect of Molecular-Scale Features on the Polymer Coil Size of Model Viscosity Index Improvers," Tribol. Lett., **62**(2), pp. 1–7.

[61] Len, M., Ramasamy, U. S., Lichter, S., and Martini, A., 2018, "Thickening Mechanisms of Polyisobutylene in Polyalphaolefin," Tribol. Lett., **66**(1).

[62] Cosimbescu, L., Vellore, A., Ramasamy, U. S., Burgess, S. A., and Martini, A., 2018, "Low Molecular Weight Polymethacrylates as Multi-Functional Lubricant Additives," Eur. Polym. J., **104**, pp. 39–44.

[63] Bhushan, B., 2000, *Modern Tribology Handbook, Two Volume Set*, CRC Press, Boca Raton.

[64] Jukic, A., Vidovic, E., and Janovic, Z., 2007, "Alkyl Methacrylate and

Styrene Terpolymers as Lubricating Oil Viscosity Index Improvers," Chem. Technol. Fuels Oils, **43**(5), pp. 386–394.

[65] Mohamad, S. A., Ahmed, N. S., Hassanein, S. M., and Rashad, A. M., 2012, "Investigation of Polyacrylates Copolymers as Lube Oil Viscosity Index Improvers," J. Pet. Sci. Eng., **100**, pp. 173–177.

[66] Šoljić Jerbić, I., Parlov Vuković, J., and Jukić, A., 2012, "Production and Application Properties of Dispersive Viscosity Index Improvers," Ind. Eng. Chem. Res., **51**(37), pp. 11914–11923.

[67] Wang, J., Ye, Z., and Zhu, S., 2007, "Topology-Engineered Hyperbranched High-Molecular-Weight Polyethylenes as Lubricant Viscosity-Index Improvers of High Shear Stability," pp. 1174–1178.

[68] Morgan, S., Ye, Z., Subramanian, R., and Zhu, S., 2010, "Higher-Molecular-Weight Hyperbranched Polyethylenes Containing Crosslinking Structures as Lubricant Viscosity-Index Improvers," Polym. Eng. Sci., **50**(5), pp. 911–918.

[69] Nunez, C. M., Chiou, B. Sen, Andrady, A. L., and Khan, S. A., 2000, "Solution Rheology of Hyperbranched Polyesters and Their Blends with Linear Polymers," Macromolecules, **33**(5), pp. 1720–1726.

[70] Knothe, G., and Steidley, K. R., 2005, "Kinematic Viscosity of Biodiesel Fuel Components and Related Compounds. Influence of Compound Structure and Comparison to Petrodiesel Fuel Components," Fuel, **84**(9), pp. 1059–1065.

[71] Gaciño, F. M., Regueira, T., Lugo, L., Comuñas, M. J. P., and Fernández, J., 2011, "Influence of Molecular Structure on Densities and Viscosities of Several Ionic Liquids," J. Chem. Eng. Data, **56**(12), pp. 4984–4999.

[72] Bhattacharya, P., Ramasamy, U. S., Krueger, S., Robinson, J. W., Tarasevich, B. J., Martini, A., and Cosimbescu, L., 2016, "Trends in Thermoresponsive Behavior of Lipophilic Polymers," Ind. Eng. Chem. Res., **55**(51), pp. 12983–12990.

[73] Van Ravensteijn, B. G. P., Bou Zerdan, R., Seo, D., Cadirov, N., Watanabe, T., Gerbec, J. A., Hawker, C. J., Israelachvili, J. N., and Helgeson, M. E., 2019, "Triple Function Lubricant Additives Based on Organic-Inorganic Hybrid Star Polymers: Friction Reduction, Wear Protection, and Viscosity Modification," ACS Appl. Mater. Interfaces, **11**(1), pp. 1363–1375.

[74] Kioupis, L. I., and Maginn, E. J., 1999, "Molecular Simulation of Poly-α-

Olefin Synthetic Lubricants: Impact of Molecular Architecture on Performance Properties," J. Phys. Chem. B, **103**(49), pp. 10781–10790.

[75] Khabaz, F., and Khare, R., 2014, "Effect of Chain Architecture on the Size, Shape, and Intrinsic Viscosity of Chains in Polymer Solutions: A Molecular Simulation Study," J. Chem. Phys., **141**(21).

[76] Ramasamy, U. S., Lichter, S., and Martini, A., 2016, "Effect of Molecular-Scale Features on the Polymer Coil Size of Model Viscosity Index Improvers," Tribol. Lett., **62**(2).

[77] Ramasamy, U. S., Len, M., and Martini, A., 2017, "Correlating Molecular Structure to the Behavior of Linear Styrene–Butadiene Viscosity Modifiers," Tribol. Lett., **65**(4).

[78] Panwar, P., Michael, P., Devlin, M., and Martini, A., 2020, "Critical Shear Rate of Polymer-Enhanced Hydraulic Fluids," Lubricants, **8**(12), pp. 1–15.

[79] Panwar, P., Len, M., Gajghate, N., Michael, P., and Martini, A., 2020, "Fluid Effects on Mechanical Efficiency of Hydraulic Pumps: Dynamometer Measurements and Molecular Simulations," ASME/BATH 2019 Symp. Fluid Power Motion Control. FPMC 2019.

[80] Selby, T. W., 1958, "The Non-Newtonian Characteristics of Lubricating Oils," ASLE Trans.

[81] Berry, G. C., 1988, "Remarks on a Relation among the Intrinsic Viscosity, the Radius of Gyration, and the Translational Friction Coefficient," J. Polym. Sci. Part B Polym. Phys., **26**(5), pp. 1137–1142.

[82] Farmer, B. S., Terao, K., and Mays, J. W., 2006, "Characterization of Model Branched Polymers by Multi-Detector SEC in Good and Theta Solvents," Int. J. Polym. Anal. Charact., **11**(1), pp. 3–19.

[83] Nassar, A. M., 2008, "Synthesis and Evaluation of Viscosity Index Improvers and Pour Point Depressant for Lube Oil," Pet. Sci. Technol., **26**(5), pp. 523–531.

[84] Ghosh, P., and Das, M., 2014, "Study of the Influence of Some Polymeric Additives as Viscosity Index Improvers and Pour Point Depressants - Synthesis and Characterization," J. Pet. Sci. Eng., **119**, pp. 79–84.

[85] Mary, C., Philippon, D., Lafarge, L., Laurent, D., Rondelez, F., Bair, S., and Vergne, P., 2013, "New Insight into the Relationship between Molecular Effects and the Rheological Behavior of Polymer-Thickened Lubricants under High Pressure," Tribol. Lett.

[86] LaRiviere, D., Asfour, A. F. A., Hage, A., and Gao, J. Z., 2000, "Viscometric Properties of Viscosity Index Improvers in Lubricant Base Oil over a Wide Temperature Range. Part I: Group II Base Oil," Lubr. Sci., **12**(2), pp. 133–143.

[87] Mueller, H. G., 1978, "Mechanism of Action of Viscosity Index Improvers," Tribol. Int., (June), pp. 189–192.

[88] Radke, W., and Müller, A. H. E., 2005, "Synthesis and Characterization of Comb-Shaped Polymers by SEC with on-Line Light Scattering and Viscometry Detection," Macromolecules, **38**(9), pp. 3949–3960.

[89] Lashkhi, V. L., and Fuks, I. G., 1988, "Polymer-Compounded Oils. Properties and Specific Features of Application," Chem. Technol. Fuels Oils, **24**(11), pp. 492–498.

[90] Holtzinger, J., Green, J., Lamb, G., Atkinson, D., and Spikes, H., 2012, "Peer-Reviewed: New Method of Measuring Permanent Viscosity Loss of Polymer-Containing Lubricants," Tribol. Lubr. Technol., **68**(12), pp. 631–639.

[91] Covitch, M. J., and Trickett, K. J., 2015, "How Polymers Behave as Viscosity Index Improvers in Lubricating Oils," Adv. Chem. Eng. Sci., **05**(02), pp. 134–151.

[92] Physique, L., Faure, D., Elf, C., France, A., and Recherche, D., "The Effect of Mechanical Stress on Four Viscosity Improver Polymers," **00**(August 2000).

[93] Stability, S., 2000, "Influence of Viscosity Index Improver , Molecular Weight , and Base Oil On," **00**(May).

[94] Campbell, K. B., Erck, R., Swita, M., and Cosimbescu, L., 2020, "Multifunctional Tunable Polymethacrylates for Enhanced Shear Stability and Wear Prevention," ACS Appl. Polym. Mater., **2**(7), pp. 2839–2848.

[95] Fenske, M., Klaus, E., and Dannenbrink, R., 1951, "Viscosity-Shear Behavior of Two Non-Newtonian Polymer-Blended Oils," *In Proceedings of the Symposium on Methods of Measuring Viscosity at High Rates of Shear, Washington, DC, USA, 24 February 1950*, ASTM International: West Conshohocken, PA, USA, 1951.

[96] ASTM D5621-07, 2013, "Standard Test Method for Sonic Shear Stability of Hydraulic Fluids," ASTM Int. West Conshohocken, PA.

[97] Coordinating European Council, 2014, *Viscosity Shear Stability of*

Transmission Lubricants (Taper Roller Bearing Rig), Coordinating European Council: Brussels, Belgium.

[98] Cui, S. T., Cummings, P. T., Cochran, H. D., Moore, J. D., and Gupta, S. A., 1998, "Nonequilibrium Molecular Dynamics Simulation of the Rheology of Linear and Branched Alkanes," Int. J. Thermophys., **19**(2 SPEC.ISS.), pp. 449–459.

[99] Bair, S., McCabe, C., McCabe, C., Cummings, P. T., and Cummings, P. T., 2002, "Comparison of Nonequilibrium Molecular Dynamics with Experimental Measurements in the Nonlinear Shear-Thinning Regime," Phys. Rev. Lett., **88**(5), pp. 583021–583024.

[100] Cecil, R., Pike, W. C., and Raje, N. R., 1973, "Development of Methods for Evaluating Traction Fluids," Wear, **26**(3), pp. 335–353.

[101] Haseltine, M. W., Duling, I. N., Hagstrom, P. E., Stenger, R. J., and Gates, D. S., 1971, "Design and Development of Fluids for Traction and Friction Type Transmissions," SAE Tech. Pap.

[102] Muraki, M., 1987, "Molecular Structure of Synthetic Hydrocarbon Oils and Their Rheological Properties Governing Traction Characteristics," Tribol. Int., **20**(6), pp. 347–354.

[103] Kyotani, T., Yoshitake, H., Ito, T., and Tamai, Y., 1986, "Correlation between Flow Properties and Traction of Lubricating Oils," ASLE Trans., **29**(1), pp. 102–106.

[104] Tsubouchi, T., and Hata, H., 1996, "Quantitative Correlations between the Fundamental Molecular Structures and Traction Properties of Traction Fluids," Japanese J. Tribol., **41**(5), pp. 463–474.

[105] Hata, H., and Tsubouchi, T., 1998, "Molecular Structures of Traction Fluids in Relation to Traction Properties," Tribol. Lett., **5**(1), pp. 69–74.

[106] Tsubouchi, T., and Hata, H., 1995, "Study on the Fundamental Molecular Structures of Synthetic Traction Fluids: Part 2," Tribol. Int., **28**(5), pp. 335–340.

[107] Dare-Edwards, M. P., 1991, "A Novel Family of Traction Fluids Deriving from Molecular Design," J. Synth. Lubr., **8**(3), pp. 197–205.

[108] Tsubouchi, T., Abe, K., and Hata, H., 1993, "Quantitative Correlation between Molecular Structures of Traction Fluids and Their Traction Properties (Part 1): Influence of Alkylene Chain," Japanese J. Tribol., **38**(3), pp. 403–410.

[109] Tsubouchi, T., and Shinoda, J., 2009, "Characterization of Oily High Bulk Modulus Fluid," World Tribol. Congr. 2009 - Proc., p. 532.

[110] Yamano, H., Shiota, K., Miura, R., Katagiri, M., Kubo, M., Stirling, A., Broclawik, E., Miyamoto, A., and Tsubouchi, T., 1996, "Molecular Dynamics Simulation of Traction Fluid Molecules under EHL Condition," Thin Solid Films, **281–282**(1–2), pp. 598–601.

[111] Washizu, H., Sanda, S., Hyodo, S., Ohmori, T., Nishino, N., and Suzuki, A., 2007, "A Molecular Dynamics Analysis of the Traction Fluids," SAE Tech. Pap.

[112] Lu, J., Wang, Q. J., Ren, N., and Lockwood, F. E., 2019, "Correlation between Pressure-Viscosity Coefficient and Traction Coefficient of the Base Stocks in Traction Lubricants: A Molecular Dynamic Approach," Tribol. Int., **134**, pp. 328–334.

[113] Washizu, H., Ohmori, T., and Suzuki, A., 2017, "Molecular Origin of Limiting Shear Stress of Elastohydrodynamic Lubrication Oil Film Studied by Molecular Dynamics," arXiv.

[114] Gattinoni, C., Heyes, D. M., Lorenz, C. D., and Dini, D., 2013, "Traction and Nonequilibrium Phase Behavior of Confined Sheared Liquids at High Pressure," Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys., **88**(5).

[115] ASTM D445-18, 2010, "Standard Test Method for Kinematic Viscosity of Transparent and Opaque Liquids (and Calculation of Dynamic Viscosity)," Annu. B. ASTM Stand.

[116] ASTM D40052-15, 2013, "Standard Test Method for Density , Relative Density , and API Gravity of Liquids by Digital Density Meter," ASTM Int., pp. 1–8.

[117] ASTM D92, 2012, "Standard Test Method for Flash and Fire Points by Cleveland Open Cup Tester," Annu. B. ASTM Stand.

[118] ISO 4409, 2007, "Hydraulic Fluid Power -- Positive-Displacement Pumps, Motors and IntegralTtransmissions -- Methods of Testing and Presenting Basic Steady State Performance," Int. Organ. Stand.

[119] Grandelli, H. E., Dickmann, J. S., Devlin, M. T., Hassler, J. C., and Kiran, E., 2013, "Volumetric Properties and Internal Pressure of Poly(α-Olefin) Base Oils," Ind. Eng. Chem. Res., **52**(50), pp. 17725–17734.

[120] Rach, S. F., and Kühn, F. E., 2009, "On the Way to Improve the Environmental Benignity of Chemical Processes: Novel Catalysts for a Polymerization Process," Sustainability, **1**(1), pp. 35–42.

[121] Plimpton, S., 1995, "Fast Parallel Algorithms for Short-Range Molecular Dynamics," J. Comput. Phys., **117**(1), pp. 1–19.

[122] Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J., 1996, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," J. Am. Chem. Soc., **118**(45), pp. 11225–11236.

[123] Siu, S. W. I., Pluhackova, K., and Böckmann, R. A., 2012, "Optimization of the OPLS-AA Force Field for Long Hydrocarbons.," J. Chem. Theory Comput., **8**(4), pp. 1459–70.

[124] Ewen, J. P., Gattinoni, C., Thakkar, F. M., Morgan, N., Spikes, H. A., and Dini, D., 2016, "A Comparison of Classical Force-Fields for Molecular Dynamics Simulations of Lubricants," Materials (Basel)., **9**(8), pp. 1–17.

[125] NosÉ, S., 2002, "A Molecular Dynamics Method for Simulations in the Canonical Ensemble," Mol. Phys., **100**(1), pp. 191–198.

[126] Hoover, W. G., 1985, "Canonical Dynamics: Equilibrium Phase-Space Distributions," Phys. Rev. A, **31**(3), pp. 1695–1697.

[127] Evans, D. J., and Morriss, G. P., 1984, "Nonlinear-Response Theory for Steady Planar Couette Flow," Phys. Rev. A, **30**(3), pp. 1528–1530.

[128] Daivis, P. J., and Todd, B. D., 2006, "A Simple, Direct Derivation and Proof of the Validity of the SLLOD Equations of Motion for Generalized Homogeneous Flows," J. Chem. Phys., **124**(19).

[129] Todd, B. D., and Daivis, P. J., 2017, *Nonequilibrium Molecular Dynamics: Theory, Algorithms and Applications*, Cambridge University Press.

[130] Ewen, J. P., Heyes, D. M., and Dini, D., 2018, "Advances in Nonequilibrium Molecular Dynamics Simulations of Lubricants and Additives," Friction, **6**(4), pp. 349–386.

[131] Panwar, P., and Michael, P., 2018, "Empirical Modelling of Hydraulic Pumps and Motors Based upon the Latin Hypercube Sampling Method," Int. J. Hydromechatronics, **1**(3), p. 272.

[132] ISO 4392-1, 2002, "Hydraulic Fluid Power - Determination of Characteristics of Motors - Part 1: At Constant Low Speed and Constant Pressure," Int. Organ. Stand.

[133] Seeton, C. J., 2006, "Viscosity–Temperature Correlation for Liquids," Tribol. Lett., **22**(1), pp. 67–78.

[134] Ulbrich, N., 2013, "Regression Model Optimization for the Analysis of Experimental Data," *In 47th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition*, p. 1344.

[135] Baker, J., and Ivantysynova, M., 2008, "Investigation Of Power Losses In The Lubricating Gap Between The Cylinder Block And Valve Plate Of Axial Piston Machines," *Proceedings of the 5th FPNI PhD Symposium*, pp. 302–319.

[136] Ivantysynova, M., and Baker, J., 2009, "Power Loss in the Lubricating Gap between Cylinder Block and Valve Plate of Swash Plate Type Axial Piston Machines," Int. J. Fluid Power, **10**(2), pp. 29–43.

[137] Morrison, F., 2001, "Understanding Rheology," Oxford Univ. Press, pp. 231–232.

[138] Kremer, K., and Grest, G. S., 1990, "Dynamics of Entangled Linear Polymer Melts: A Molecular-Dynamics Simulation," J. Chem. Phys., **92**(8), pp. 5057–5086.

[139] Mondello, M., and Grest, G. S., 2002, "Viscosity Calculations of n - Alkanes by Equilibrium Molecular Dynamics Viscosity Calculations of n-Alkanes by Equilibrium Molecular Dynamics," **9327**(March 1997).

[140] Mondello, M., Grest, G. S., Webb, E. B., and Peczak, P., 1998, "Dynamics of N-Alkanes: Comparison to Rouse Model," J. Chem. Phys., **109**(2), pp. 798–805.

[141] Shimida, T., Horng, P. L., Porter, R. S., Shimida, T., Horng, P. L., and Roger, S., 1980, "Broad Shear Range Viscometry of High Polymer Solutions : Polystyrene and Polyisobutene in Decalin Broad Shear Range Viscometry of High Polymer Solutions : Polystyrene and Polyisobutene in Decalin," **783**.

[142] Zolper, T. J., Seyam, A. M., Chen, C., Jungk, M., Stammer, A., Stoegbauer, H., Marks, T. J., Chung, Y. W., and Wang, Q., 2013, "Energy Efficient Siloxane Lubricants Utilizing Temporary Shear-Thinning," Tribol. Lett., **49**(3), pp. 525–538.

[143] Tanner, R. I., 1985, *Engineering Rheology.*, Oxford University Press: Oxford, UK.

[144] Rouse, P. E., 1953, "A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers," J. Chem. Phys., **21**(7), pp. 1272–1280.

[145] Rouse, P. E., and Sittel, K., 1953, "Viscoelastic Properties of Dilute

Polymer Solutions," J. Appl. Phys., **24**(6), pp. 690–696.

[146] De Gennes, P. G., 1971, "Reptation of a Polymer Chain in the Presence of Fixed Obstacles," J. Chem. Phys., **55**(2), pp. 572–579.

[147] De Gennes, P. G., and Gennes, P. G., 1979, *Scaling Concepts in Polymer Physics*, Cornell university press.

[148] de Gennes, P. G., 1980, "Dynamics of Fluctuations and Spinodal Decomposition in Polymer Blends," J. Chem. Phys., **72**(9), pp. 4756–4763.

[149] Doi, M., and Edwards, S. F., 1988, *The Theory of Polymer Dynamics*, oxford university press.

[150] Mondello, M., Grest, G. S., Garcia, A. R., Silbernagel, B. G., Mondello, M., Grest, G. S., Garcia, A. R., and Silbernagel, B. G., 1998, "Molecular Dynamics of Linear and Branched Alkanes : Simulations and Nuclear Magnetic Resonance Results Nuclear Magnetic Resonance Results," **5208**(June 1996).

[151] Robbins, M. O., 2017, "Probing Large Viscosities in Glass-Formers with Nonequilibrium Simulations," **114**(39), pp. 7952–7957.

[152] Jadhao, V., and Robbins, M. O., 2019, "Rheological Properties of Liquids Under Conditions of Elastohydrodynamic Lubrication," Tribol. Lett.

[153] Martin, M. G., and Siepmann, J. I., 1998, "Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes," J. Phys. Chem. B, **102**(14), pp. 2569–2577.

[154] Martin, M. G., and Siepmann, J. I., 1999, "Novel Configurational-Bias Monte Carlo Method for Branched Molecules. Transferable Potentials for Phase Equilibria. 2. United-Atom Description of Branched Alkanes," J. Phys. Chem. B, **103**(21), pp. 4508–4517.

[155] Wick, C. D., Martin, M. G., and Siepmann, J. I., 2000, "Transferable Potentials for Phase Equilibria. 4. United-Atom Description of Linear and Branched Alkenes and Alkylbenzenes," J. Phys. Chem. B, **104**(33), pp. 8008–8016.

[156] Dinpajooh, M., and Nitzan, A., 2020, "Heat Conduction in Polymer Chains with Controlled End-to-End Distance," J. Chem. Phys., **153**(16).

[157] Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and Mackerell, A. D., 2010, "CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological

Force Fields," J. Comput. Chem., **31**(4), pp. 671–690.

[158] Reiher, W. E., 1985, "Theoretical Studies of Hydrogen Bonding," Harvard University.

[159] Maginn, E. J., Messerly, R. A., Carlson, D. J., Roe, D. R., and Elliott, J. R., 2019, "Best Practices for Computing Transport Properties 1. Self-Diffusivity and Viscosity from Equilibrium Molecular Dynamics [Article v1.0]," Living J. Comput. Mol. Sci., **1**(1), pp. 1–20.

[160] Zhang, Y., Otani, A., and Maginn, E. J., 2015, "Reliable Viscosity Calculation from Equilibrium Molecular Dynamics Simulations: A Time Decomposition Method," J. Chem. Theory Comput., **11**(8), pp. 3537–3546.

[161] Zhmud, B., 2014, "Viscosity Blending Equations," Lube, **121**, pp. 24–29.

[162] Fetters, L. J., Graessley, W. W., and Kiss, A. D., 1991, "Viscoelastic Properties of Polyisobutylene Melts," (3), pp. 3136–3141.

[163] Ferry, J. D., 1980, *Viscoelastic Properties of Polymers*, John Wiley & Sons: Hoboken, NJ, USA.

[164] Bair, S., and Winer, W. O., 2007, "A Quantitative Test of the Einstein-Debye Relation Using the Shear Dependence of Viscosity for Low Molecular Weight Liquids," Tribol. Lett., **26**(3), pp. 223–228.

[165] Bird, R B. Curtiss, C. F., 1987, *Dynamics of Polymeric Fluids Volume 2: Kinetic Theory*, Wiley: Hoboken, NJ, USA.

[166] Carreau, P. J., 1972, "Rheological Equations from Molecular Network Theories," Trans. Soc. Rheol., **16**(1), pp. 99–127.

[167] Nichetti, D., and Manas-Zloczower, I., 1998, "Viscosity Model for Polydisperse Polymer Melts," J. Rheol. (N. Y. N. Y)., **42**(4), pp. 951–969.

[168] Brown, S. F., 2015, "Base Oil Groups: Manufacture, Properties, and Performance," Tribol. Lubr. Technol., **71**(4), pp. 32–35.

[169] Gartner, T. E., and Jayaraman, A., 2019, "Modeling and Simulations of Polymers: A Roadmap," Macromolecules, **52**(3), pp. 755–786.

[170] Yeh, I. C., and Hummer, G., 2004, "System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions," J. Phys. Chem. B, **108**(40), pp. 15873–15879.

[171] Landau, L. D., Lifshitz, E. M., and Reichl, L. E., 1981, "Statistical Physics, Part 1," Phys. Today, **31**(1).

[172] Robbins, M. O., Grest, G. S., and Kremer, K., 1990, "Effect of Finite System Size on Thermal Fluctuations: Implications for Melting," Phys. Rev. B, **42**(9), pp. 5579–5585.

[173] Mondello, M., Grest, G. S., Mondello, M., and Grest, G. S., 1998, "Molecular Dynamics of Linear and Branched Alkanes Molecular Dynamics of Linear and Branched Alkanes," **7156**(June 1995).

[174] Yang, T., Sun, Y., Meng, X., Wu, J., and Ilja Siepmann, J., 2021, "Simultaneous Measurement of the Density and Viscosity for N-Decane + CO2 Binary Mixtures at Temperature between (303.15 to 373.15) K and Pressures up to 80 MPa," J. Mol. Liq., **338**.

[175] Chen, B., Potoff, J. J., and Siepmann, J. I., 2002, "Monte Carlo Calculations for Alcohols and Their Mixtures with Alkanes. Transferable Potentials for Phase Equilibria. 5. United-Atom Description of Primary, Secondary, and Tertiary Alcohols," J. Phys. Chem. B, **105**(15), pp. 3093–3104.

[176] Stubbs, J. M., Potoff, J. J., and Siepmann, J. I., 2004, "Transferable Potentials for Phase Equilibria. 6. United-Atom Description for Ethers, Glycols, Ketones, and Aldehydes," J. Phys. Chem. B, **108**(45), pp. 17596–17605.

[177] Kamath, G., Cao, F., and Potoff, J. J., 2004, "An Improved Force Field for the Prediction of the Vapor-Liquid Equilibria for Carboxylic Acids," J. Phys. Chem. B, **108**(37), pp. 14130–14136.

[178] Kamath, G., Robinson, J., and Potoff, J. J., 2006, "Application of TraPPE-UA Force Field for Determination of Vapor-Liquid Equilibria of Carboxylate Esters," Fluid Phase Equilib., **240**(1), pp. 46–55.

[179] Maerzke, K. A., Schultz, N. E., Ross, R. B., and Siepmann, J. I., 2009, "TraPPE-UA Force Field for Acrylates and Monte Carlo Simulations for Their Mixtures with Alkanes and Alcohols," J. Phys. Chem. B, **113**(18), pp. 6415–6425.

[180] Yang, Y., Narayanan Nair, A. K., and Sun, S., 2020, "Sorption and Diffusion of Methane and Carbon Dioxide in Amorphous Poly(Alkyl Acrylates): A Molecular Simulation Study," J. Phys. Chem. B, **124**(7), pp. 1301–1310.

[181] Ferrando, N., Gedik, I., Lachet, V., Pigeon, L., and Lugo, R., 2013, "Prediction of Phase Equilibrium and Hydration Free Energy of Carboxylic Acids by Monte Carlo Simulations," J. Phys. Chem. B, **117**(23), pp. 7123–7132.

[182] Clifford, S., Bolton, K., and Ramjugernath, D., 2006, "Monte Carlo Simulation of Carboxylic Acid Phase Equilibria," J. Phys. Chem. B, **110**(43), pp. 21938–21943.

[183] Washizu, H., Sanda, S., Hyodo, S. A., Ohmori, T., Nishino, N., and Suzuki, A., 2008, "All-Atom Molecular Dynamics Simulation of Submicron Thickness EHL Oil Film," 2007 Proc. ASME/STLE Int. Jt. Tribol. Conf. IJTC 2007, **PART A**, pp. 167–169.

[184] Maćkowiak, S., Heyes, D. M., Dini, D., and Brańka, A. C., 2016, "Non-Equilibrium Phase Behavior and Friction of Confined Molecular Films under Shear: A Non-Equilibrium Molecular Dynamics Study," J. Chem. Phys., **145**(16).

[185] Han, X. J., and Schober, H. R., 2011, "Transport Properties and Stokes-Einstein Relation in a Computer-Simulated Glass-Forming Cu33.3Zr66.7 Melt," Phys. Rev. B - Condens. Matter Mater. Phys., **83**(22).

[186] Lerner, E., and Bouchbinder, E., 2017, "Effect of Instantaneous and Continuous Quenches on the Density of Vibrational Modes in Model Glasses," Phys. Rev. E, **96**(2).

[187] Shi, J., Wang, J., Yi, X., and Fan, X., 2021, "Effect of Film Thickness on Slip and Traction Performances in Elastohydrodynamic Lubrication by a Molecular Dynamics Simulation," Tribol. Lett., **69**(4).

[188] Bas G. P. van Ravensteijn, Raghida Bou Zerdan, Craig J. Hawker, and M. E. H., 2021, "Role of Architecture on Thermorheological Properties of Poly(Alkyl Methacrylate)-Based Polymers," Macromolecules, **54**(12), pp. 5473–5483.

[189] Mattice, W. L., and Suter, U. W., 1994, *Conformational Theory of Large Molecules*, Wiley-Interscience, New York.

[190] Theodorou, D. N., and Suter, U. W., 1985, "Shape of Unperturbed Linear Polymers: Polypropylene," Macromolecules, **18**(6), pp. 1206–1214.

[191] Cané, E., Llovell, F., and Vega, L. F., 2017, "Accurate Viscosity Predictions of Linear Polymers from N-Alkanes Data," J. Mol. Liq., **243**, pp. 115–123.

[192] Chávez Thielemann, H., Cardellini, A., Fasano, M., Bergamasco, L., Alberghini, M., Ciorra, G., Chiavazzo, E., and Asinari, P., 2019, "From GROMACS to LAMMPS: GRO2LAM: A Converter for Molecular Dynamics Software," J. Mol. Model., **25**(6).

[193] Plimpton, S., and others, 2007, "LAMMPS-Large-Scale Atomic/Molecular Massively Parallel Simulator," Sandia Natl. Lab.

[194] David, A., Pearlman, and, David, A., Case, and, James, and W., 1995, "AMBER, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules," Comput. Phys. Commun.

[195] Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., Van Der Spoel, D., Hess, B., and Lindahl, E., 2013, "GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit," Bioinformatics, **29**(7), pp. 845–854.

[196] Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M., 2009, "CHARMM: The Biomolecular Simulation Program," J. Comput. Chem., **30**(10), pp. 1545–1614.

[197] Shan, Y., Kim, E. T., Eastwood, M. P., Dror, R. O., Seeliger, M. A., and Shaw, D. E., 2011, "How Does a Drug Molecule Find Its Target Binding Site?," J. Am. Chem. Soc., **133**(24), pp. 9181–9183.

[198] BIOVIA, and Dassault Systèmes, 2020, "Materials Studio," San Diego Dassault Systèmes.

[199] Phillips, J. C., Schulten, K., Bhatele, A., Mei, C., Sun, Y., Bohm, E. J., and Kale, L. V., 2016, "Scalable Molecular Dynamics with NAMD," Parallel Sci. Eng. Appl. Charm++ Approach, pp. 60–76.

[200] Smidstrup, S., Markussen, T., Vancraeyveld, P., Wellendorff, J., Schneider, J., Gunst, T., Verstichel, B., Stradi, D., Khomyakov, P. A., Vej-Hansen, U. G., Lee, M. E., Chill, S. T., Rasmussen, F., Penazzi, G., Corsetti, F., Ojanperä, A., Jensen, K., Palsgaard, M. L. N., Martinez, U., Blom, A., Brandbyge, M., and Stokbro, K., 2020, "QuantumATK: An Integrated Platform of Electronic and Atomic-Scale Modelling Tools," J. Phys. Condens. Matter, **32**(1).

[201] Dias, C. S., 2021, "Molecular Dynamics Simulations of Active Matter Using LAMMPS."

[202] Moriwaki, H., Tian, Y. S., Kawashita, N., and Takagi, T., 2018, "Mordred: A Molecular Descriptor Calculator," J. Cheminform., **10**(1).

[203] Dong, J., Cao, D. S., Miao, H. Y., Liu, S., Deng, B. C., Yun, Y. H., Wang, N. N., Lu, A. P., Zeng, W. Bin, and Chen, A. F., 2015, "ChemDes: An Integrated Web-Based Platform for Molecular Descriptor and Fingerprint Computation," J. Cheminform., **7**(1).

[204] Grover, M., Singh, B., Bakshi, M., and Singh, S., 2000, "Quantitative Structure-Property Relationships in Pharmaceutical Research - Part 2," Pharm. Sci. Technol. Today, **3**(2), pp. 50–57.

[205] Grover, M., Singh, B., Bakshi, M., and Singh, S., 2000, "Quantitative Structure-Property Relationships in Pharmaceutical Research - Part 1," Pharm. Sci. Technol. Today, **3**(2), pp. 50–57.

[206] Selassie, C. D., Garg, R., Kapur, S., Kurup, A., Verma, R. P., Mekapati, S. B., and Hansch, C., 2002, "Comparative QSAR and the Radical Toxicity of Various Functional Groups," Chem. Rev., **102**(7), pp. 2585–2605.

[207] Katritzky, A. R., Kuanar, M., Slavov, S., Hall, C. D., Karelson, M., Kahn, I., and Dobchev, D. A., 2010, "Quantitative Correlation of Physical and Chemical Properties with Chemical Structure: Utility for Prediction," Chem. Rev., **110**(10), pp. 5714–5789.

[208] Dearden John, and Worth Andrew, 2007, *In Silico Prediction of Physicochemical Properties*, Luxembourg.

[209] Dearden, J. C., Rotureau, P., and Fayet, G., 2013, "QSPR Prediction of Physico-Chemical Properties for REACH," SAR QSAR Environ. Res., **24**(4), pp. 279–318.

[210] Mauri, A., Consonni, V., Pavan, M., and Todeschini, R., 2006, "DRAGON Software: An Easy Approach to Molecular Descriptor Calculations," Match, **56**(2), pp. 237–248.

[211] Todeschini, R., and Consonni, V., 2000, *Handbook of Molecular Descriptors*.

[212] Yap C. W., 2010, "PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints," J. Comput. Chem., **32**(7), pp. 1466–1474.

[213] Georg, H., 2008, "Blue-Desc Descriptor Calculator," Tübingen, Ger. Univ. Tübingen.

[214] Landrum, G., 2011, "RDKit : A Software Suite for Cheminformatics , Computational Chemistry , and Predictive Modeling," Components.

[215] Steinbeck, C., Han, Y., Kuhn, S., Horlacher, O., Luttmann, E., and Willighagen, E., 2003, "The Chemistry Development Kit (CDK): An Open-Source Java Library for Chemo- and Bioinformatics.," ChemInform, **34**(21).

[216] O'Boyle, N. M., and Hutchison, G. R., 2008, "Cinfony - Combining Open Source Cheminformatics Toolkits behind a Common Interface," Chem. Cent. J., **2**(1).

[217] Cao, D. S., Xu, Q. S., Hu, Q. N., and Liang, Y. Z., 2013, "ChemoPy: Freely Available Python Package for Computational Biology and Chemoinformatics," Bioinformatics, **29**(8), pp. 1092–1094.

[218] Holland, R. C. G., Down, T. A., Pocock, M., Prlić, A., Huen, D., James, K., Foisy, S., Dräger, A., Yates, A., Heuer, M., and Schreiber, M. J., 2008, "BioJava: An Open-Source Framework for Bioinformatics," Bioinformatics, **24**(18), pp. 2096–2097.

[219] Dong, J., Yao, Z. J., Wen, M., Zhu, M. F., Wang, N. N., Miao, H. Y., Lu, A. P., Zeng, W. Bin, and Cao, D. S., 2016, "BioTriangle: A Web-Accessible Platform for Generating Various Molecular Representations for Chemicals, Proteins, DNAs/RNAs and Their Interactions," J. Cheminform., **8**(1).

[220] Spjuth, O., Helmus, T., Willighagen, E. L., Kuhn, S., Eklund, M., Wagener, J., Murray-Rust, P., Steinbeck, C., and Wikberg, J. E. S., 2007, "Bioclipse: An Open Source Workbench for Chemo- and Bioinformatics," BMC Bioinformatics, **8**.

[221] Cao, D. S., Xu, Q. S., and Liang, Y. Z., 2013, "Propy: A Tool to Generate Various Modes of Chou's PseAAC," Bioinformatics, **29**(7), pp. 960–962.

[222] Cao, D. S., Liang, Y. Z., Yan, J., Tan, G. S., Xu, Q. S., and Liu, S., 2013, "PyDPI: Freely Available Python Package for Chemoinformatics, Bioinformatics, and Chemogenomics Studies," J. Chem. Inf. Model., **53**(11), pp. 3086–3096.

[223] Liu, B., Liu, F., Fang, L., Wang, X., and Chou, K. C., 2015, "RepDNA: A Python Package to Generate Various Modes of Feature Vectors for DNA Sequences by Incorporating User-Defined Physicochemical Properties and Sequence-Order Effects," Bioinformatics, **31**(8), pp. 1307–1309.

[224] Kuhn, T., Willighagen, E. L., Zielesny, A., and Steinbeck, C., 2010, "CDK-

Taverna: An Open Workflow Environment for Cheminformatics," BMC Bioinformatics, **11**.

[225] Xiao, N., Cao, D. S., Zhu, M. F., and Xu, Q. S., 2015, "Protr/ProtrWeb: R Package and Web Server for Generating Various Numerical Representation Schemes of Protein Sequences," Bioinformatics, **31**(11), pp. 1857–1859.

[226] Hinselmann, G., Rosenbaum, L., Jahn, A., Fechner, N., and Zell, A., 2011, "JCompoundMapper: An Open Source Java Library and Command-Line Tool for Chemical Fingerprints," J. Cheminform., **3**(1).

[227] Cao, Y., Charisi, A., Cheng, L. C., Jiang, T., and Girke, T., 2008, "ChemmineR: A Compound Mining Framework for R," Bioinformatics, **24**(15), pp. 1733–1734.

[228] Cao, D. S., Xiao, N., Xu, Q. S., and Chen, A. F., 2015, "Rcpi: R/Bioconductor Package to Generate Various Descriptors of Proteins, Compounds and Their Interactions," Bioinformatics, **31**(2), pp. 279–281.

[229] Weininger, D., 1988, "SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules," J. Chem. Inf. Comput. Sci., **28**(1), pp. 31–36.

[230] Zhang, B., Van Duin, A. C. T., and Johnson, J. K., 2014, "Development of a ReaxFF Reactive Force Field for Tetrabutylphosphonium Glycinate/CO2 Mixtures," J. Phys. Chem. B, **118**(41), pp. 12008–12016.

[231] Panwar, P., Schweissinger, E., Maier, S., Hilf, S., Sirak, S., and Martini, A., 2022, "Effect of Polymer Structure and Chemistry on Viscosity Index, Thickening Efficiency, and Traction Coefficient of Lubricants," J. Mol. Liq., **359**.

[232] Erhan, S. Z., and Asadauskas, S., 2000, "Lubricant Basestocks from Vegetable Oils," Ind. Crops Prod., **11**(2–3), pp. 277–282.

[233] Jameel, A. G. A., and Sarathy, S. M., 2006, "Lube Products: MolecularCharacterization of Base Oils," Encycl. Anal. Chem. Appl. Theory Instrum., pp. 1–14.

[234] Höglund, E., 1999, "Influence of Lubricant Properties on Elastohydrodynamic Lubrication," Wear, **232**(2), pp. 176–184.

[235] Moity, L., Molinier, V., Benazzouz, A., Barone, R., Marion, P., and Aubry, J. M., 2014, "In Silico Design of Bio-Based Commodity Chemicals: Application to Itaconic Acid Based Solvents," Green Chem., **16**(1), pp. 146–160.

[236] Nieto-Draghi, C., Fayet, G., Creton, B., Rozanska, X., Rotureau, P., De Hemptinne, J. C., Ungerer, P., Rousseau, B., and Adamo, C., 2015, "A General Guidebook for the Theoretical Prediction of Physicochemical Properties of Chemicals for Regulatory Purposes," Chem. Rev., **115**(24), pp. 13093–13164.

[237] Roy, K., Kar, S., and Das, R. N., 2015, "A Primer on QSAR/QSPR Modeling: Fundamental Concepts," SpringerBriefs Mol. Sci., **1**, pp. 1–121.

[238] Nantasenamat, C., Isarankura-Na-Ayudhya, C., Naenna, T., and Prachayasittikul, V., 2009, "A Practical Overview of Quantitative Structure-Activity Relationship," EXCLI J., **8**, pp. 74–88.

[239] Ajmani, S., Jadhav, K., and Kulkarni, S. A., 2009, "Group-Based QSAR (G-QSAR): Mitigating Interpretation Challenges in QSAR," QSAR Comb. Sci., **28**(1), pp. 36–51.

[240] Oskarsdottir, G., Bhan, A., Venkatasubramanian, V., Thomson, K. T., Snively, C. M., Katare, S., Lauterbach, J. A., and Caruthers, J. M., 2003, "Catalyst Design: Knowledge Extraction from High-Throughput Experimentation," J. Catal., **216**(1–2), pp. 98–109.

[241] Bernazzani, L., Duce, C., Micheli, A., Mollica, V., Sperduti, A., Starita, A., and Tiné, M. R., 2006, "Predicting Physical-Chemical Properties of Compounds from Molecular Structures by Recursive Neural Networks," J. Chem. Inf. Model., **46**(5), pp. 2030–2042.

[242] Ivanciuc, O., Ivanciuc, T., Filip, P. A., and Cabrol-Bass, D., 1999, "Estimation of the Liquid Viscosity of Organic Compounds with a Quantitative Structure-Property Model," J. Chem. Inf. Comput. Sci., **39**(3), pp. 515–524.

[243] Suzuki, T., Ohtaguchi, K., and Koide, K., 1996, "Computer-Assisted Approach to Develop a New Prediction Method of Liquid Viscosity of Organic Compounds," Comput. Chem. Eng., **20**(2), pp. 161–173.

[244] Suzuki, T., Ebert, R. U., and Schüürmann, G., 1997, "Development of Both Linear and Nonlinear Methods to Predict the Liquid Viscosity at 20 °C of Organic Compounds," J. Chem. Inf. Comput. Sci., **37**(6), pp. 1122–1128.

[245] Suzuki, T., Ebert, R. U., and Schüürmann, G., 2001, "Application of Neural Networks to Modeling and Estimating Temperature-Dependent Liquid Viscosity of Organic Compounds," J. Chem. Inf. Comput. Sci., **41**(3), pp. 776–790.

[246] Katritzky, A. R., Chen, K., Wang, Y., Karelson, M., Lucic, B., Trinajstic, N., Suzuki, T., and Schüürmann, G., 2000, "Prediction of Liquid Viscosity for Organic Compounds by a Quantitative Structure-Property Relationship," J. Phys. Org. Chem., **13**(1), pp. 80–86.

[247] Lučić, B., Bašic, I., Nadramija, D., Miličević, A., Trinajstić, N., Suzuki, T., Petrukhin, R., Karelson, M., and Katritzky, A. R., 2002, "Correlation of Liquid Viscosity with Molecular Structure for Organiccompounds Using Different Variable Selection Methods," Arkivoc, **2002**(4), pp. 45–59.

[248] Cocchi, M., De Benedetti, P. G., Seeber, R., Tassi, L., and Ulrici, A., 1999, "Development of Quantitative Structure-Property Relationships Using Calculated Descriptors for the Prediction of the Physicochemical Properties (ND, ρ, Bp, ∈, η) of a Series of Organic Solvents," J. Chem. Inf. Comput. Sci., **39**(6), pp. 1190–1203.

[249] Kauffman, G. W., and Jurs, P. C., 2001, "Prediction of Surface Tension, Viscosity, and Thermal Conductivity for Common Organic Solvents Using Quantitative Structure-Property Relationships," J. Chem. Inf. Comput. Sci., **41**(2), pp. 408–418.

[250] Koutsoukos, S., Philippi, F., Malaret, F., and Welton, T., 2021, "A Review on Machine Learning Algorithms for the Ionic Liquid Chemical Space," Chem. Sci., **12**(20), pp. 6820–6843.

[251] Loh, G. C., Lee, H. C., Tee, X. Y., Chow, P. S., and Zheng, J. W., 2020, "Viscosity Prediction of Lubricants by a General Feed-Forward Neural Network," J. Chem. Inf. Model., **60**(3), pp. 1224–1234.

[252] Goussard, V., Duprat, F., Ploix, J. L., Dreyfus, G., Nardello-Rataj, V., and Aubry, J. M., 2020, "A New Machine-Learning Tool for Fast Estimation of Liquid Viscosity. Application to Cosmetic Oils," J. Chem. Inf. Model., **60**(4), pp. 2012–2023.

[253] Nashawi, I. S., and Elgibaly, A. A., 1999, "Prediction of Liquid Viscosity of Pure Organic Compounds via Artificial Neural Networks," Pet. Sci. Technol., **17**(9), pp. 1107–1144.

[254] Heintz, J., Belaud, J. P., Pandya, N., Teles Dos Santos, M., and Gerbaud, V., 2014, "Computer Aided Product Design Tool for Sustainable Product Development," Comput. Chem. Eng., **71**, pp. 362–376.

[255] Dreyfus, G., 2005, *Neural Networks: Methodology and Applications*.

[256] Goulon, A., Picot, T., Duprat, A., and Dreyfus, G., 2007, "Predicting

Activities without Computing Descriptors: Graph Machines for QSAR," SAR QSAR Environ. Res., **18**(1–2), pp. 141–153.

[257] Petitjean, M., 1992, "Applications of the Radius-Diameter Diagram to the Classification of Topological and Geometrical Shapes of Chemical Compounds," J. Chem. Inf. Comput. Sci., **32**(4), pp. 331–337.

[258] Dehmer, M., Varmuza, K., and Bonchev, D., 2012, *Statistical Modelling of Molecular Descriptors in QSAR/QSPR*.

[259] Todeschini, R., and Consonni, V., 2010, *Molecular Descriptors for Chemoinformatics*.

[260] JERZY LESZCZYNSKI, 2010, "Recent Advances in QSAR Studies," Igarss 2014, (8), p. 428.

[261] Gordon, P. A., 2003, "Characterizing Isoparaffin Transport Properties with Stokes-Einstein Relationships," Ind. Eng. Chem. Res., **42**(26), pp. 7025–7036.

[262] Panwar, P., Yang, Q., and Martini, A., 2022, "PyL3dMD: Python LAMMPS 3D Molecular Descriptors Package," J. Chem. Inf. Model., **XXX**(XXXX), pp. XX–XX.

[263] American Petroleum Institute, 1967, *Properties of Hydrocarbons of High Molecular Weight Synthesized by Research Project 42 of the American Petroleum Institute*, New York.

[264] Martinez, L., Andrade, R., Birgin, E. G., and Martínez, J. M., 2009, "PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations," J. Comput. Chem., **30**(13), pp. 2157–2164.

[265] Dodda, L. S., De Vaca, I. C., Tirado-Rives, J., and Jorgensen, W. L., 2017, "LigParGen Web Server: An Automatic OPLS-AA Parameter Generator for Organic Ligands," Nucleic Acids Res., **45**(W1), pp. W331–W336.

[266] Jorgensen, W. L., 1998, "BOSS - Biochemical and Organic Simulation System," Encycl. Comput. Chem., **3**, pp. 3281–3285.

[267] Dodda, L. S., Vilseck, J. Z., Tirado-Rives, J., and Jorgensen, W. L., 2017, "1.14∗CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations," J. Phys. Chem. B, **121**(15), pp. 3864–3870.

[268] Elisseeff, A., and Guyon, I., 2003, "An Introduction to Variable and Feature Selection.," J. Mach. Learn. Res., pp. 1157–1182.

[269] Tibshirani, R., 1996, "Regression Shrinkage and Selection Via the

Lasso," J. R. Stat. Soc. Ser. B, **58**(1), pp. 267–288.

[270] Zou, H., and Hastie, T., 2005, "Regularization and Variable Selection via the Elastic Net," J. R. Stat. Soc. Ser. B Stat. Methodol., **67**(2), pp. 301–320.

[271] Miller, R. E., 1985, "Analysis of Variance.," Chem. Eng. (New York), **92**(6), pp. 173–178.

[272] Belsley, D. A., E. Kuh, and R. E. W., 1980, *Regression Diagnostics*, John Wiley & Sons, Inc., New York, NY.

[273] Gogtay, N. J., and Thatte, U. M., 2017, "Principles of Correlation Analysis," J. Assoc. Physicians India, **65**(MARCH), pp. 78–81.

[274] Rasmussen, C. E. and C. K. I. W., 2006, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Massachusetts.

[275] Hastie, T., Tibshirani, R., and Friedman, J., 2001, *The Elements of Statistical Learning*, Springer New York, New York, NY.

[276] O'Brien, R. M., 2007, "A Caution Regarding Rules of Thumb for Variance Inflation Factors," Qual. Quant., **41**(5), pp. 673–690.

[277] Snoek, J., Larochelle, H., and Adams, R. P., 2012, "Practical Bayesian Optimization of Machine Learning Algorithms," Adv. Neural Inf. Process. Syst., **4**, pp. 2951–2959.

[278] Gelbart, M. A., Snoek, J., and Adams, R. P., 2014, "Bayesian Optimization with Unknown Constraints," Uncertain. Artif. Intell. - Proc. 30th Conf. UAI 2014, pp. 250–259.

[279] Bull, A. D., 2011, "Convergence Rates of Efficient Global Optimization Algorithms," J. Mach. Learn. Res., **12**, pp. 2879–2904.

[280] Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and De Freitas, N., 2016, "Taking the Human out of the Loop: A Review of Bayesian Optimization," Proc. IEEE, **104**(1), pp. 148–175.

[281] jerome h, friedman, 2001, "Greedy Function Approximation a Gradient Boosting Machine," Ann. Stat., pp. 1189–1232.

[282] Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E., 2015, "Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation," J. Comput. Graph. Stat., **24**(1), pp. 44–65.

[283] Ribeiro, M. T., Singh, S., and Guestrin, C., 2016, "'Why Should i Trust

You?' Explaining the Predictions of Any Classifier," Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., **13-17-Augu**, pp. 1135–1144.

[284] Wright, R., 2018, "Interpreting Black-Box Machine Learning Models Using Partial Dependence and Individual Conditional Expectation Plots," Explor. SAS ® Enterp. Min. Spec. Collect., pp. 1950–2018.

[285] Broto,, P; Devilers, J., 1990, "Autocorrelation of Properties Distributed on Molecular Graphs."

[286] Sanderson, R. T., 1983, "Electronegativity and Bond Energy," J. Am. Chem. Soc., **105**(8), pp. 2259–2261.

[287] Wang, S. Q., 2007, "On Chain Statistics and Entanglement of Flexible Linear Polymer Melts," Macromolecules, **40**(24), pp. 8684–8694.

[288] Hall, L. H., and Kier, L. B., 2007, "The Molecular Connectivity Chi Indexes and Kappa Shape Indexes in Structure-Property Modeling," Rev. Comput. Chem., **2**, pp. 367–422.

[289] Colby, R. H., Fetters, L. J., and Graessley, W. W., 1987, "Melt Viscosity-Molecular Weight Relationship for Linear Polymers," Macromolecules, **20**(9), pp. 2226–2237.

[290] Morgan, M. J., Mukwembi, S., and Swart, H. C., 2012, "A Lower Bound on the Eccentric Connectivity Index of a Graph," Discret. Appl. Math., **160**(3), pp. 248–258.

[291] Pettersson, A., 2007, "High-Performance Base Fluids for Environmentally Adapted Lubricants," Tribol. Int., **40**(4), pp. 638–645.

[292] Gedde, U. W., 1999, *Polymer Physics*, Springer Netherlands, Berlin.

[293] Chen, Y., Jha, S., Raut, A., Zhang, W., and Liang, H., 2020, "Performance Characteristics of Lubricants in Electric and Hybrid Vehicles: A Review of Current and Future Needs," Front. Mech. Eng., **6**.

[294] Callen, H. B., and Scott, H. L., 1998, "Thermodynamics and an Introduction to Thermostatistics, 2nd Ed.," Am. J. Phys., **66**(2), pp. 164–167.

[295] Kwak, Y., Cleveland, C., Adhvaryu, A., Fang, X., Hurley, S., and Adachi, T., 2019, "Understanding Base Oils and Lubricants for Electric Drivetrain Applications," SAE Tech. Pap., (December).

[296] Gantzer, P., Creton, B., and Nieto-Draghi, C., 2020, "Inverse-QSPR for de Novo Design: A Review," Mol. Inform., **39**(4).

[297] Hasko, D., Shang, L., Noppe, E., and Lefrançois, E., 2019, "Virtual Assessment and Experimental Validation of Power Loss Contributions in Swash Plate Type Axial Piston Pumps," Energies, **12**(16).

[298] Reiher III, W. H., 1985, "Theoretical Studies of Hydrogen Bonding," Harvard University, Cambridge, MA, USA.

[299] PCS Instruments, 2014, "MTM2 Mini-Traction Machine" [Online]. Available: https://pcs-instruments.com/wp-content/uploads/2014/03/MTM2.pdf.

[300] Morgado, P. L., Echávarri Otero, J., Sánchez-Peñuela Lejarraga, J. B., Muñoz Sanz, J. L., Díaz Lantada, A., Muñoz-Guijosa, J. M., Lorenzo Yustos, H., and Leal Wiña, P., 2009, "Models for Predicting Friction Coefficient and Parameters with Influence in Elastohydrodynamic Lubrication," Adv. Tribol., pp. 144–145.

Appendix A.

A.1. Instruments

A Cannon StressTech HR Oscillatory Rheometer was used to measure dynamic viscosity at low shear rates and the PCS Ultra Shear viscometer was used to measure the dynamic viscosity at high shear conditions. Schematics of these devices are shown in Figure A.1.

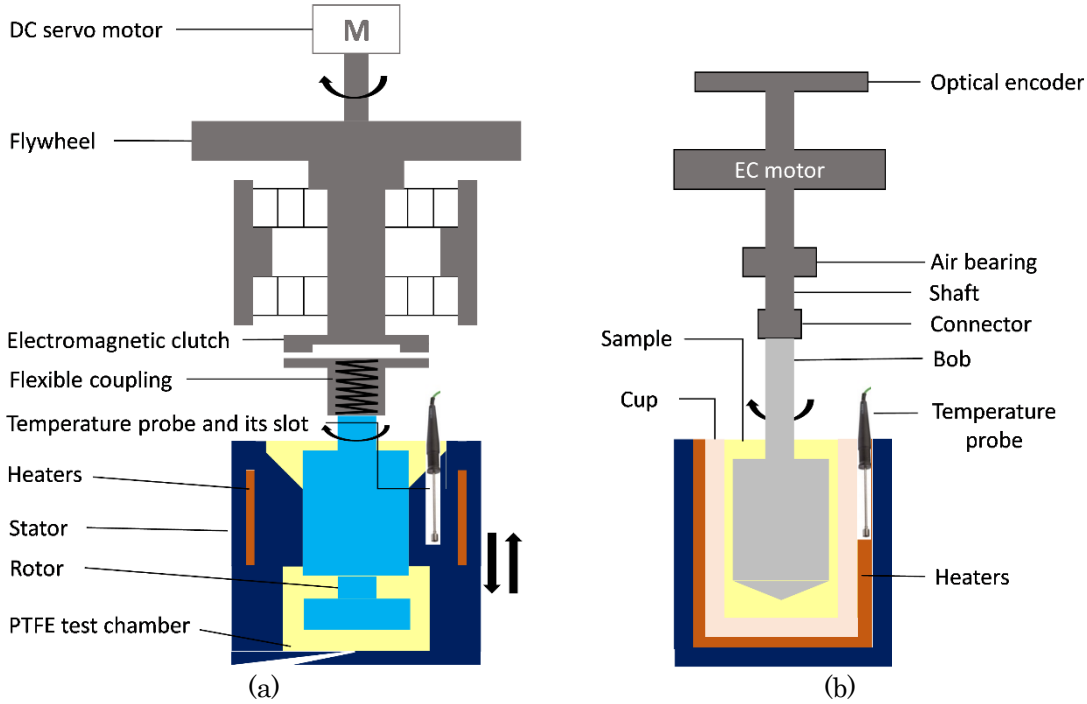


Figure A.1 Schematics of (a) PCS Ultra Shear Viscometer and (b) Cannon StressTech HR Oscillatory Rheometer used for measuring high and low shear viscosity, respectively.

A.2. Force Field Parameters

The TraPPE-UA forcefield was used to describe the chemistry of molecules. In the TraPPE-UA force field, the nonbonded interactions are described by pairwise-additive Lennard-Jones 12-6 potentials, electrostatic interactions by the Ewald summation, bonded atoms are considered to have fixed bond lengths, bond angles are governed by harmonic potential, and the motion of the dihedral angle is governed by Fourier potential. The original TraPPE uses fixed bond lengths, but to model fully flexible bonds, a harmonic potential was implemented by taking the corresponding force constants from the CHARMM [157,298] force field, as suggested by Siepmann [153]. This fully flexible model can be described by Equation A.1.

$$U_{total} = U_{bond}(r) + U_{bend}(\theta) + U_{torsion}(\phi) + U_{NB}(r_{ij})$$

$$= \sum_{bond} \frac{k_r}{2}(r - r_0)^2 + \sum_{torsion} \sum_{i=1}^{m} C_i[1 + cos(n_i\phi - d_i)] \quad \text{(A.1)}$$

$$+ \sum_{bend} \frac{k_\theta}{2}(\theta - \theta_0)^2 + \sum_{i<j} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

Here, r is the bond length, r_0 is the equilibrium bond length, k_r is the bond length force constant, θ is the bond angle, θ_0 is the equilibrium bond angle, k_θ is the bond angle force constant, r_{ij} is the site-site separation, ϵ_{ij} is the LJ well depth and diameter and q_i and q_j are the partial charges on sites i and j, respectively. For LJ interactions between two different atom types, the standard Lorentz-Bethelot empirical combining rules were used to compute parameters using Equation A.2.

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad \text{and} \quad \epsilon_{ij} = (\epsilon_{ii}\epsilon_{jj})^{1/2} \quad \text{(A.2)}$$

For the saturated and unsaturated hydrocarbons, all pseudo-atoms were connected to pseudo-atoms formed from heavy atoms of the same electronegativity; thus, partial charges are not needed here for the TraPPE-UA potential.

The potential parameters that are used to fully describe these molecules are given in Tables A.1, A.2, A.3, and A.4.

Table A.1 UA force field parameters for non-bonded interactions [153–155].

Pseudo/United atom	ϵ/k_B [K]	σ [Å]	q [e]
CH_3 (sp^3)	98.0	3.750	0.0
CH_2 (sp^3)	46.0	3.950	0.0
CH_2 (sp^2)	85.0	3.675	0.0
CH (sp^3)	10.0	4.680	0.0
C (sp^3)	0.5	6.400	0.0
C (sp^2)	20.0	3.850	0.0

Table A.2 UA force field parameters for 1-2 bonded interactions [153–157,298].

Bond	r_0 [Å]	k_r/k_B [K/Å^2]
CH_x-CH_y	1.54	452900
CH_x=CH_y	1.33	825280

Table A.3 UA force field parameters for 1-3 bonded interactions [153–155].

Angle	r_θ [°]	k_θ/k_B [K/Å^2]
CH_x-CH_2-CH_y	114.0	62500
CH_x-CH-CH_y	112.0	62500
CH_x-C-CH_y	109.5	62500
CH_x=C-CH_y	119.7	70420

Table A.4 UA force field parameters for 1-4 bonded interactions [153–155].

Torsion	c_0/k_B [K]	c_1/k_B [K]	c_2/k_B [K]	c_3/k_B [K]
CH_x-CH_2-CH_2-CH_y	0.00	355.03	-68.19	791.32
CH_x-CH_2-CH-CH_y	-251.06	428.73	-111.85	441.27
CH_x-CH_2-C-CH_y	0.00	0.00	0.00	461.29
CH_x-CH_2-C=CH_y	688.50	86.36	-109.77	-282.24

A.3. Viscosity Simulation Details

The low-shear or Newtonian viscosity of fluids was modeled using the Green-Kubo (GK) approach which relate the shear viscosity to the integral over time of the pressure tensor autocorrelation function [159,160].

$$\eta(t) = \frac{V}{6k_BT}\int_0^\infty \langle P_{\alpha\beta}(t)\cdot P_{\alpha\beta}(0)\rangle dt \tag{A.3}$$

Here, V is the system volume, k_B is the Boltzmann constant, T is the temperature, $P_{\alpha\beta}$ denotes the element $\alpha\beta$ of the pressure or stress tensor, and the angle bracket indicates the ensemble average.

The high shear viscosity η in the non-equilibrium molecular dynamics simulation was determined by computed the ratio of shear stress P_{xz} to the shear rate $\dot{\gamma}$.

$$\eta = -\frac{P_{xz}}{\dot{\gamma}} \tag{A.4}$$

A.4. Governing Equations

Rouse Model: The Rouse model [140,145,164] provides the relaxation times as a function of density ρ and viscosity η_0 of fluid, and the molecular weight M of polymer as

$$1/\lambda = \frac{\pi^2}{12}\frac{\rho R_g T}{\eta_0 M} \tag{A.5}$$

Bird et al. Model: This model [11,165] provides the relaxation times as a function of density ρ and viscosity η_0 of fluid, and the molecular weight M and concentration c_p of polymer. In addition to this, it considers the viscosity η_{0s} of solvent which is base oil in this case.

$$1/\lambda = \frac{c_p \rho R_g T}{(\eta_0 - \eta_{0s})M} \tag{A.6}$$

Kendall-Monroe: The kinematic viscosity of blend of polymer and base oil was determined by the Kendall-Monroe equation which calculate the viscosity of a blend as the cubic-root average of the viscosity of its components,

$$\mu^{1/3} = x_p \mu_p^{1/3} + x_s \mu_s^{1/3} \tag{A.7}$$

Here, μ, μ_p, and μ_s are the kinematic viscosity of blend, polymer, and base oil or solvent, x_p and x_s are the mole fraction of polymer and base oil in the blend, respectively.

William–Landel–Ferry (WLF): The Newtonian viscosity of polyisobutylene (PIB) as a function of its molecular weight was calculated using an empirical model. This empirical model has been extracted from a large set of experimental data to provide viscosity in mPa·s at 25 °C [162,163].

$$\eta_{0p} = 4.69 \times 10^{-9} M^{3.43} \tag{A.8}$$

The viscosity of PIB at any temperature T can be obtained by a temperature shift factor which is described by the WLF equation as

$$\log(a_T) = \log\left[\eta_{0p}(T)/\eta_{0p}(T_0)\right] = -\frac{c_1(T - T_0)}{c_2 + (T - T_0)} \tag{A.9}$$

Here, T_0 is the reference temperature of 25 °C and empirical parameters c_1 = 8.61 and c_2 = 200 K [162,163].

A.5. Carreau Model

Parameter of the Carreau fit to experiment, and simulation viscosity data are tabulated in Table A.5.

Table A.5 Parameter of the Carreau fit to experiment and simulation viscosity data.

Temperature	Fluid ID	Parameters			
		η_0 [mPa·s]	η_∞ [mPa·s]	λ [s]	n [-]
50 °C	Fluid 1	40.149	0.000	1.898	0.763

	Fluid 2	61.917	1.564	3.592	0.712
	Fluid 3	35.451	0.000	3.693	0.671
80 °C	Fluid 1	15.837	0.000	4.143	0.793
	Fluid 2	22.26	0.000	7.742	0.790
	Fluid 3	13.483	0.000	9.932	0.662

Appendix B.

B.1. Instruments

A Cannon StressTech HR Oscillatory Rheometer was used to measure dynamic viscosity at low shear rates, a generic viscometer was used to measure the kinematic viscosity, and a PCS Mini-Traction Machine was used to measure traction coefficients. Schematics of these devices are shown in Figure B.1.

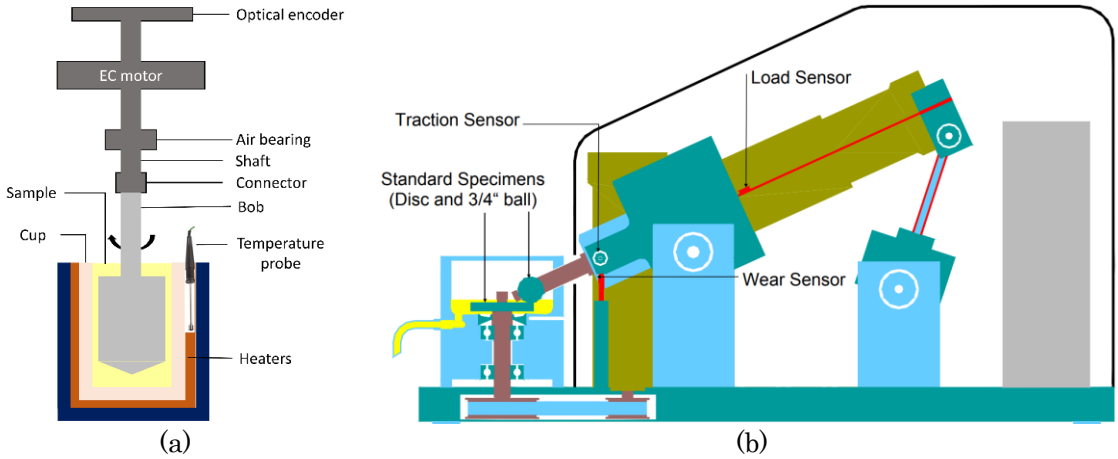


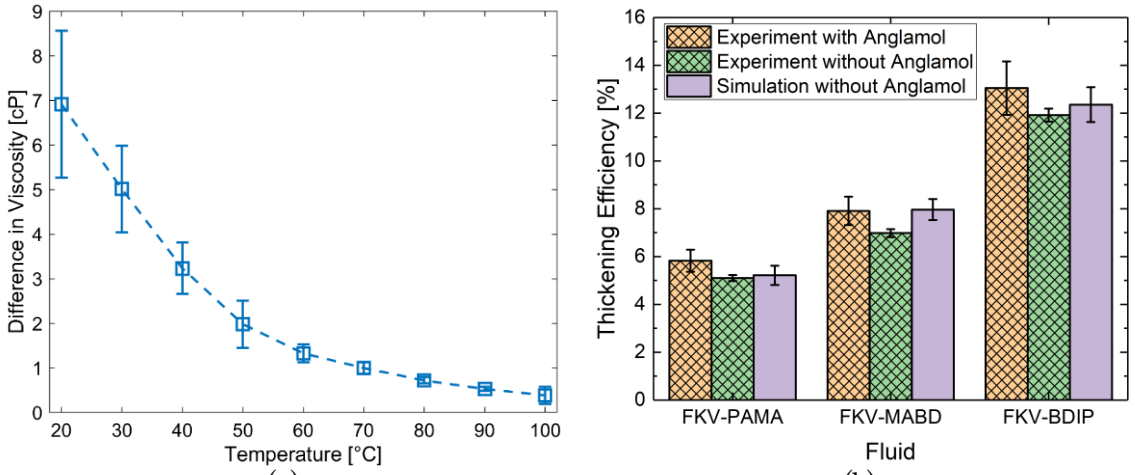
Figure B.1 Schematics of (a) Cannon StressTech HR Oscillatory Rheometer and (b) PCS Mini-Traction Machine [299] used for measuring dynamic viscosity and traction coefficient, respectively.

B.2. Experimental and Simulation Data

Listed in Table B.1 are the experimental and simulated kinematic viscosity of all fluid with error at 40 and 100 °C. The asterisk indicates the kinematic viscosity measured for the fluids with additive package. The viscosity data listed in Table B.1 and shown in Figure B.2(a) show that the fluids formulated with 6.5 wt.% of Anglamol 99 additive package had higher viscosity than the fluids formulated without the additive package.

Table B.1 Experimental and simulated viscosity of base oil and test fluids. The asterisk indicates the kinematic viscosity measured for the fluids with 6.5 wt.% of Anglamol 99 additive package.

Temperature [°C]	Fluid	Polymer, C [wt.%]	Viscosity [cSt]		
			Experiment*	Experiment	Simulation
40	Base Oil (PAO3cSt)	0.0	-	11.50 ± 0.31	11.74 ± 1.18
	FKV-PAO	15.7	34.08	-	34.31 ± 4.30
	FKV-PIB	14.6	35.60	-	33.25 ± 4.20
	FKV-BDIP	9.8	33.23 ± 0.39	30.81	28.64 ± 4.54
	FKV-PAMA	20.8	31.83 ± 0.33	28.69	27.61 ± 5.21
	FKV-MABD	15.8	32.78 ± 0.41	29.51	33.30 ± 4.81
100	Base Oil (PAO3cSt)	0.0	-	3.09 ± 0.18	3.14 ± 0.27
	FKV-PAO*	15.7	7.01	-	7.05 ± 0.80
	FKV-PIB	14.6	7.03	-	6.89 ± 0.53
	FKV-BDIP	9.8	7.06 ± 0.32	6.72	6.85 ± 0.62
	FKV-PAMA	20.8	6.86 ± 0.29	6.39	6.44 ± 0.77
	FKV-MABD	15.8	6.98 ± 0.28	6.52	7.00 ± 0.59



(a) (b)

Figure B.2 (a) Average increase in viscosity of the fluids when the fluids are formulated with 6.5 wt.% of Anglamol 99 additive package compared to the fluids formulated without the additive package. (b) Thickening efficiency calculated from experimentally measured viscosity of the fluids formulated with and without 6.5 wt.% of Anglamol additives package at 100 °C, and from the simulation calculated viscosity of the fluids formulated without 6.5 wt.% of Anglamol additives package at 100 °C.

This comparison in Figure B.2(b) shows a consistent trend in thickening efficiency of fluids formulated with and without 6.5 wt.% of Anglamol additives package. Listed in Table B.2 are the experimental and simulated traction coefficients of all fluid at 40 °C and MTM test conditions. The asterisk indicates the traction coefficient measured for the fluids with additive package. The traction coefficient data listed in Table B.2 shows that the fluids formulated

with 6.5 wt.% of Anglamol 99 additive package had higher traction coefficient than the fluids formulated without the additive package.

Table B.2 Experimental and simulated traction coefficient of test fluids. The asterisk indicates the traction coefficient measured for the fluids with 6.5 wt.% of Anglamol 99 additive package.

Fluid	Polymer, C [wt.%]	Traction Coefficient [-]		
		Experiment*	Experiment	Simulation
FKV-PAO	15.7	0.04550	-	0.05068
FKV-PIB	14.6	0.05900	-	0.06583
FKV-BDIP	9.8	0.04650	0.04460	0.05672
FKV-PAMA	20.8	0.04850	0.04720	0.06434
FKV-MABD	15.8	0.04780	0.04630	0.05796

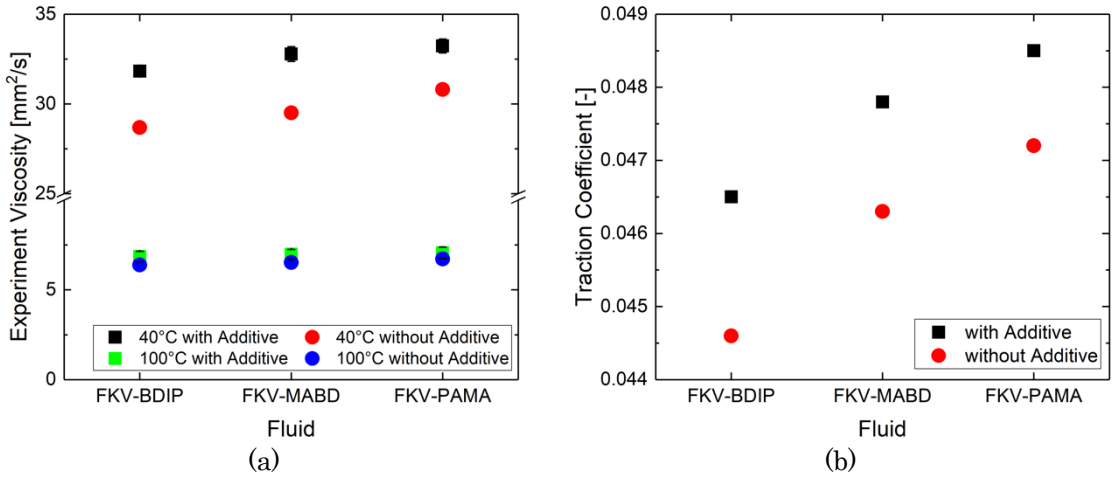


Figure B.3 (a) Experimentally measured kinematic viscosity of the fluids formulated with and without 6.5 wt.% of Anglamol additives package at 40 and 100°C. (b) Experimentally measured traction coefficient of the fluids formulated with and without 6.5 wt.% of Anglamol additives package at 40°C.

This comparison in Figure B.3 shows a consistent trend in viscosity and traction coefficient of fluids formulated with and without 6.5 wt.% of Anglamol additives package.

B.3. Base Oil Formulation

Since the composition of group III 3 cSt base oil was unknown, an alternative 3 cSt base oil was formulated by blending 25 wt.% of 2 cSt PAO with 75 wt.% of 4 cSt PAO for the simulation. Previous gas chromatography-mass spectroscopy measurements reported that PAO2 consisted of nearly 100% 1-decene dimer $C_{20}H_{42}$ and PAO4 consisted of equal parts 1-decene trimer $C_{30}H_{62}$ and tetramer $C_{40}H_{82}$ [119]. Viscosity of the blended samples was measured at low shear rate of 50 1/s (that is, in Newtonian regime) at 40 and 100 °C using

a Cannon StressTech HR Oscillatory Rheometer. The average and standard deviation are determined from 3 runs each of 2 blended samples of base oil. The properties of this 3 cSt PAO base oil are listed in Table B.3. It must be noted here that all the fluids were simulated by blending polymers with this synthetic 3 cSt PAO base oil, however, physically blended fluids with polymers used group III 3 cSt base oil. The group III and group IV (i.e., PAO based base oils) have viscosity index (VI) as high as 120 and 140, respectively. This difference in the VI might have some impact on the viscosity index of the samples.

Table B.3 Properties of 3 cSt PAO base oil blend.

Temperature [°C]	Avg. Viscosity [cP]	Std. Deviation [cP]	Density [g/cc]	Avg. Viscosity [cSt]
40	9.153	0.225	0.796	11.499
100	2.324	0.168	0.753	3.086

Table B.4 Calculation of ratio of base oil molecules in the proposed 3 cSt PAO base oil.

25wt% PAO2 + 75wt% PAO4				
Molecule	PAO2	PAO4		Sum
	Dimer	Trimer	Tetramer	
Molar mass of each molecule, M_{oi} [g/mol]	282.548	422.813	563.079	1268.440
Weight fraction of each molecule, w_i	0.250	0.375	0.375	1.000
Mole of each molecule in 100 g, $n_i = w_i/M_{oi} \times 100$	0.089	0.089	0.067	0.244
Total number of moles, $\sum, n_i$	0.244			0.244
Mole fraction of each molecule, $x_i = n_i/\sum, n_i$	0.363	0.364	0.273	1.000
Number of each molecule if total number of molecules in the simulation box is $N_T = 11$, $N_i = N_T \times x_i$	3.993	4.002	3.005	11.000

B.4. Simulation Plan

For all simulations, we tried to have around or more than 12000 UA in the simulation box of approximate dimension of 300.0Å x 30.0Å x 30.0Å. These criteria and desired formulation listed in Table B.5 which decided how many polymers we should have in the simulation box for each fluid. As it can be seen from the dimension of simulation box, orthogonal simulation boxes were considered since these models were also sheared in the traction coefficients simulation.

Table B.5 Simulation plan of each test fluids.

Fluid ID	Desired Formulation	Number of Molecules
FKV-PAO	• 15.7 wt.% PAO Polymer • 84.3 wt.% 3 cSt Base oil	• 5 PAO polymer → 15.7 wt.% • 120 Dimers → 21.1 wt.% • 120 Trimers → 31.6 wt.% • 90 Tetramers → 31.6 wt.% • Box size = 296.4Å x 30.0Å x 30.0Å
FKV-PIB	• 14.6 wt.% PIB Polymer • 85.4 wt.% 3 cSt Base oil	• 11 PIB polymer → 14.6 wt.% • 122 Dimers → 21.2 wt.% • 122 Trimers → 32.1 wt.% • 91 Tetramers → 32.1 wt.% • Box size = 252.9Å x 30.0Å x 30.0Å
FKV-BDIP	• 9.8 wt.% BDIP Polymer • 90.2 wt.% 3 cSt Base oil	• 1 BDIP polymer → 9.8 wt.% • 132 Dimers → 22.6 wt.% • 132 Trimers → 33.8 wt.% • 100 Tetramers → 33.8 wt.% • Box size = 306.5Å x 30.0Å x 30.0Å
FKV-PAMA	• 20.8 wt.% PAMA Polymer • 79.2 wt.% 3 cSt Base oil	• 1 PAMA polymer → 20.8 wt.% • 110 Dimers → 19.8 wt.% • 110 Trimers → 29.7 wt.% • 83 Tetramers → 29.7 wt.% • Box size = 289.5Å x 30.0Å x 30.0Å
FKV-MABD	• 15.8 wt.% MABD Polymer • 84.2 wt.% 3 cSt Base oil	• 2 MABD polymer → 15.8 wt.% • 105 Dimers → 21.0 wt.% • 106 Trimers → 31.6 wt.% • 79 Tetramers → 31.6 wt.% • Box size = 260.5Å x 30.0Å x 30.0Å

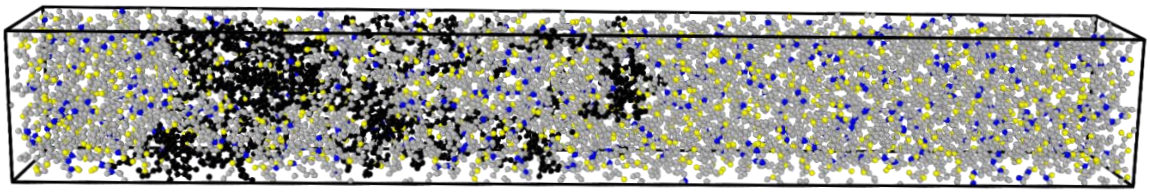


Figure B.4 Simulation box representing FKV-MABD fluid. The cluster of black atoms represents MABD polymers and cluster of yellow, grey, and blue colors atoms represent base oil molecules (dimer, trimer, and tetramer of 1-decene). The longest side of box was considered *x*-direction in which shear was applied.

B.5. Forcefield Parameters

The TraPPE-UA forcefield was used to describe the chemistry of molecules [153–156,175,177–179,181,182]. In the TraPPE-UA force field, the nonbonded

interactions U_{NB} between atoms which are separated by more than three bonds or belongs to different molecules are described by combination of pairwise-additive Lennard-Jones 12-6 potentials and electrostatic interactions by the Ewald summation, 1-2 bonded interactions U_{bond} are considered to have fixed bond lengths, 1-3 bonded or bond angles interactions U_{bend} are governed by harmonic potential, and the motion of the dihedral angle or 1-4 bonded interactions $U_{torsion}$ is governed by Fourier potential. The original TraPPE uses fixed bond lengths, but to model fully flexible bonds, a harmonic potential was implemented by taking the corresponding force constants [156,180]. This fully flexible model can be described by Equation B.1.

$$U_{total} = U_{bond}(r) + U_{bend}(\theta) + U_{torsion}(\phi) + U_{NB}(r_{ij}) \tag{B.1}$$

$$U_{bond}(r) = \frac{k_r}{2}(r - r_0)^2 \tag{B.2}$$

$$U_{bend}(\theta) = \frac{k_\theta}{2}(\theta - \theta_0)^2 \tag{B.3}$$

$$U_{torsion}(\phi) = C_0 + C_1[1 + cos(\phi)] + C_2[1 - cos(2\phi)] + C_3[1 + cos(3\phi)] + C_4[1 - cos(4\phi)] \tag{B.4}$$

$$U_{NB}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{B.5}$$

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \text{ and } \epsilon_{ij} = (\epsilon_{ii}\epsilon_{jj})^{1/2} \tag{B.6}$$

Here, U_{total} is the total interaction potential, U_{bond}, U_{bend}, $U_{torsion}$, respectively are the potential of 1-2, 1-3, and 1-4 bonded interaction, and U_{NB} is the total potential of nonbonded interactions due to both the pairwise and electrostatic interactions. r is the distance between two bonded atoms, θ is the angle between three consecutively bonded atoms, ϕ is the torsional angle between four consecutively bonded atoms. Here, r_{ij}, ϵ_{ij}, and σ_{ij} are the separation, LJ well depth, and LJ size, respectively, for the pair of atoms i and j. For Lennard-Jones interactions between two different atom types, the standard Lorentz-Bethelot empirical combining rules were used to compute parameters using Equation B.6.

It must be mentioned that TraPPE torsion potential parameters in some literatures were provided for either of the torsion potential of forms expressed by Equations B.7 and B.8.

$$U_{torsion}(\phi) = \sum_{n=0,4} A_n cos^n(\phi) \\ = A_0 + A_1 cos(\phi) + A_2 cos^2(\phi) + A_3 cos^3(\phi) + A_4 cos^4(\phi) \quad \text{(B.7)}$$

$$U_{torsion}(\phi) = B_0 + B_1 cos(\phi) + B_2 cos(2\phi) + B_3 cos(3\phi) + B_4 cos(4\phi) \quad \text{(B.8)}$$

These torsional potential parameters or coefficients were then converted using the following rigorous equivalences that were obtained by trigonometric transformations:

$C_0 = A_0 - A_1 + A_2 - A_3 + A_4$, $C_1 = A_1 + 3A_3/4$, $C_2 = -(A_2 + A_4)/2$, $C_3 = A_3/4$, and $C_4 = -A_4/8$,

or $C_0 = B_0 - B_1 + B_2 - B_3 + B_4$, $C_1 = B_1$, $C_2 = -B_2$, $C_3 = B_3$, and $C_4 = -B_4$.

The potential parameters that are used to fully describe these molecules by Equation B.11 are given in Tables B.6, B.7, B.8, and B.9.

Table B.6 Atom types and the Lennard-Jones parameters for nonbonded interactions. The color in table represents the color of atoms in the polymer molecules.

Pseudo atom	Group	Short form	Type Number	Mass	LJ potential well depth, ε/k_B [K]	LJ diameter, σ [Å]	Partial charge, q [e]
CH_3	Alkane	ch3	1	15.03450	98.0	3.75	0.00
CH_2	Alkane	ch2	2	14.02658	46.0	3.95	0.00
CH	Alkane	ch	3	13.01864	10.0	4.68	0.00
C(C=)	Alkane	c	4	12.01070	0.5	6.40	0.00
C(=O)	Carbonyl	ccbn	5	12.01070	40.0	3.82	+0.40
O(=C)	Carbonyl	ocbn	6	15.99940	79.0	3.05	-0.40
O(-CH_2)	Ether	oet	7	15.99940	55.0	2.80	-0.25
CH_2(-O-)	Ether	ch2et	8	14.02658	46.0	3.95	+0.25
C(CH-OH)	Alkane	c	9	12.01070	0.5	6.40	0.00
CH(-OH)	Alkane	chacl	10	13.01864	10.0	4.33	+0.515
O(-H)	Alcohol	oacl	11	15.99940	93.0	3.02	-0.70
H(-O)	Alcohol	hacl	12	1.007940	0.0	0.0	+0.435
O(-CH_2)	Ether	oet	13	15.99940	55.0	2.80	-0.50
CH_2(-O-)	Ether	ch2et	14	14.02658	46.0	3.95	+0.25
CH_3(-O-)	Ether	ch3et	15	15.03450	98.0	3.75	+0.25
CH(=C-)	Alkene	chsp2	16	13.01864	47.0	3.73	0.00
C(=CH-)	Alkene	csp2	17	12.01070	20.0	3.85	0.00

Table B.7 Bond parameters for 1-2 bonded interaction.

Bond	Type Number	Equilibrium bond length, r_0 [Å]	Bond energy constant, k_l/k_B [K/Å^2]
CH_x-CH_y	1	1.54	452500
CH_x-C(=O)	2	1.52	452500
C=O	3	1.20	704500
O-C(=O)	4	1.344	452500
O-CH_2	5	1.41	452500
CH_x-CH(-OH)	6	1.54	452500
CH-O(-H)	7	1.43	452500
O-H	8	0.945	704500
O-CH(-OH)	9	1.41	452500
O-CH_2	10	1.41	452500
O-CH_3	11	1.41	452500
CH_x=CH_y	12	1.33	452500

Table B.8 Angle parameters for 1-3 bonded interaction.

Angle	Type Number	Equilibrium angle, θ_0 [°]	Angle energy constant, k_θ/k_B [K/rad^2]
CH_x-CH_2-CH_y	1	114.00	62500
CH_x-CH-CH_y	2	112.00	62500
CH_x-C-CH_y	3	109.47	62500
CH_x-C-C(=O)	4	109.47	62500
CH_x-C(=O)=O	5	126.00	40300
CH_x-C(=O)-O	6	111.00	35300
O(=C)=C-O	7	123.00	40300
C(=O)-O-CH_2	8	115.00	62500
O-CH_2-CH_x	9	112.00	50300
CH_x-C-CH(-OH)	11	109.47	62500
CH_x-CH(-OH)-OH	12	109.47	50400
CH_x-CH(-OH)-O	13	112.00	50300
HO-CH-O	14	116.24	45350
CH(-OH)-O-CH_2	15	112.00	60400
O-CH_2-CH_x	16	112.00	50300
CH(-OH)-O-CH_3	17	112.00	60400
CH-O-H	18	108.50	55400
CH_x=(CH=)-CH_y	19	119.7	70420
CH_x=(C=)-CH_y	20	119.7	70420
CH_x-(C=)-CH_y	21	119.7	70420

Table B.9 Dihedral parameters for 1-4 bonded interaction.

Torsion	Type Number	Angle energy constant,			
		c_0/k_B [K]	c_1/k_B [K]	c_2/k_B [K]	c_3/k_B [K]
CH_x-(CH_2)-(CH_2)-CH_y	1	0	355.03	-68.19	791.32
CH_x-(CH_2)-(CH)-CH_y	2	-251.06	428.73	-111.85	441.27
CH_x-(CH)-(CH)-CH_y	3	-251.06	428.73	-111.85	441.27
CH_x-(CH_2)-(C)-CH_y	4	0	0	0	461.29
CH_x-(CH_2)-(C)-C(=O)	5	0	0	0	461.29
CH_x-(C)-C(=O)=O	6	2035.58	-736.9	57.84	-293.23
CH_x-(C)-C(=O)-O	7	-24.66	736.84	57.87	293.25
(C)-C(=O)-O-CH_2(-O-)	8	0.00	2158.00	2098.00	197.30
C(=O)-O- CH_2(-O-)-CH_x	9	0	725.35	-163.75	558.2
O=C(=O)-O-CH_2(-O-)	10	4715.99	-2194.00	2059.00	-153.40
O- CH_2(-O-)- CH_x-CH_y	11	0	176.62	-53.34	769.93
CH_x-(CH_2)-(C)-CH(-OH)	12	0	0	0	461.29
CH_x-(C)-CH(-OH)-OH	13	1260.00	-630.00	781.20	0.00
CH_x-(C)-CH(-OH)-O	14	-251.06	428.73	-111.85	441.27
(C)-CH(-OH)-O-H	15	215.89	197.33	31.46	-173.92
(C)-CH(-OH)-O-CH_2(-O-)	16	0	725.35	-163.75	558.20
H-O-CH(-OH)-O	17	0	630.00	781.20	0.00
HO-CH(-OH)-O-CH_2(-O-)	18	4715.99	-2194.00	2059.00	-153.40
CH(-OH)-O-CH_2(-O-)-CH_x	19	0	725.35	-163.75	558.2
HO-CH(-OH)-O-CH_3(-O-)	20	4715.99	-2194.00	2059.00	-153.40
O- CH_2(-O-)-CH_2-CH_y	21	0	176.62	-53.34	769.93
(C)-CH(-OH)-O-CH_3(-O-)	22	0	725.35	-163.75	558.20
CH_x-C-CH=C	23	688.5	86.36	-109.77	-282.24
CH_x-CH_2=C-CH_y	24	Quadratic Potential, $U_{torsion}(\phi) = d_\phi(\phi - \phi_0)^2$ d_ϕ = 12400 [K] (cis) = 13400 [K] (trans) $\phi_0 = \pi$ [rad] (cis) = 0 [rad] (trans)			

B.6. Optimization of Simulation Parameters for Forcefield

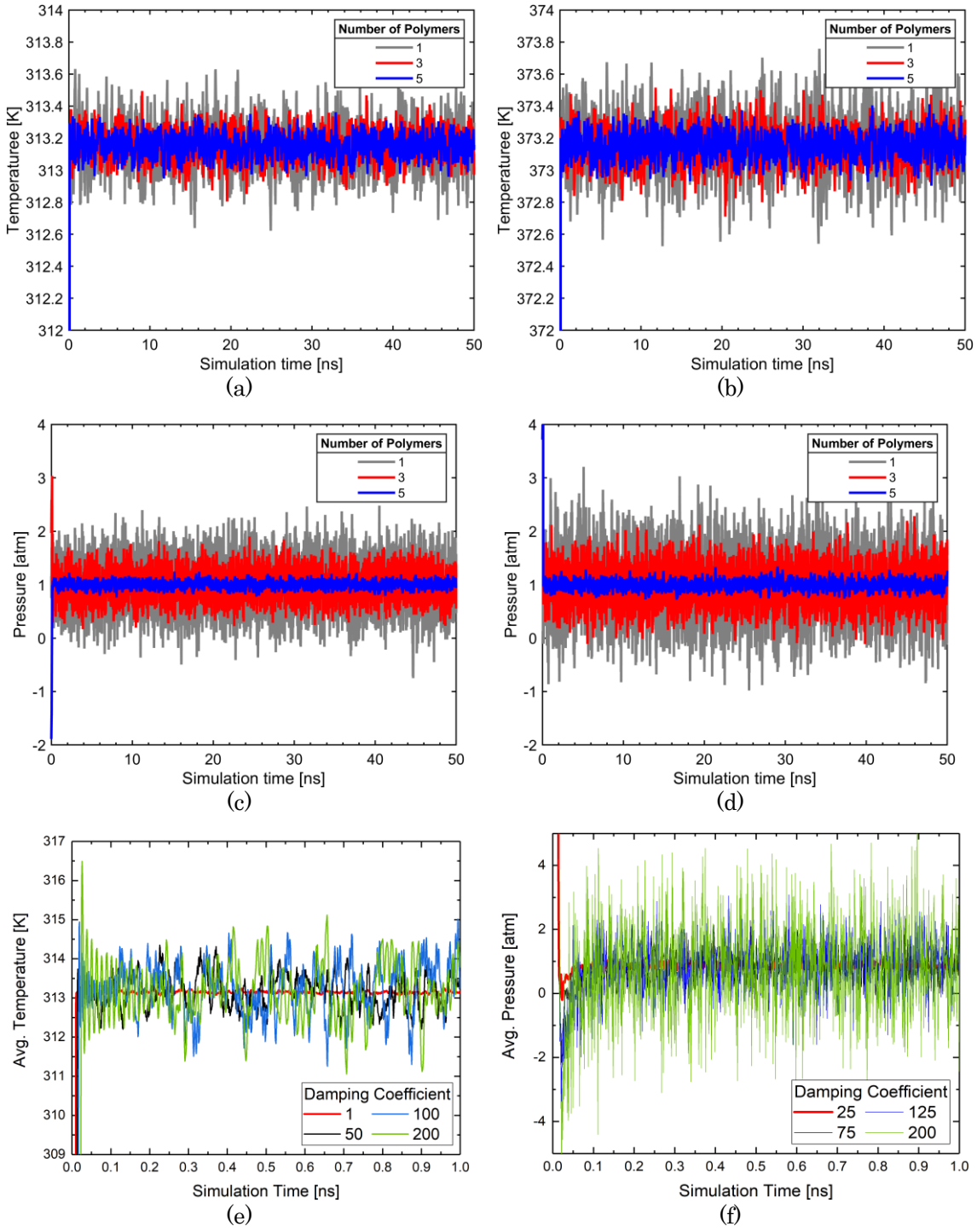
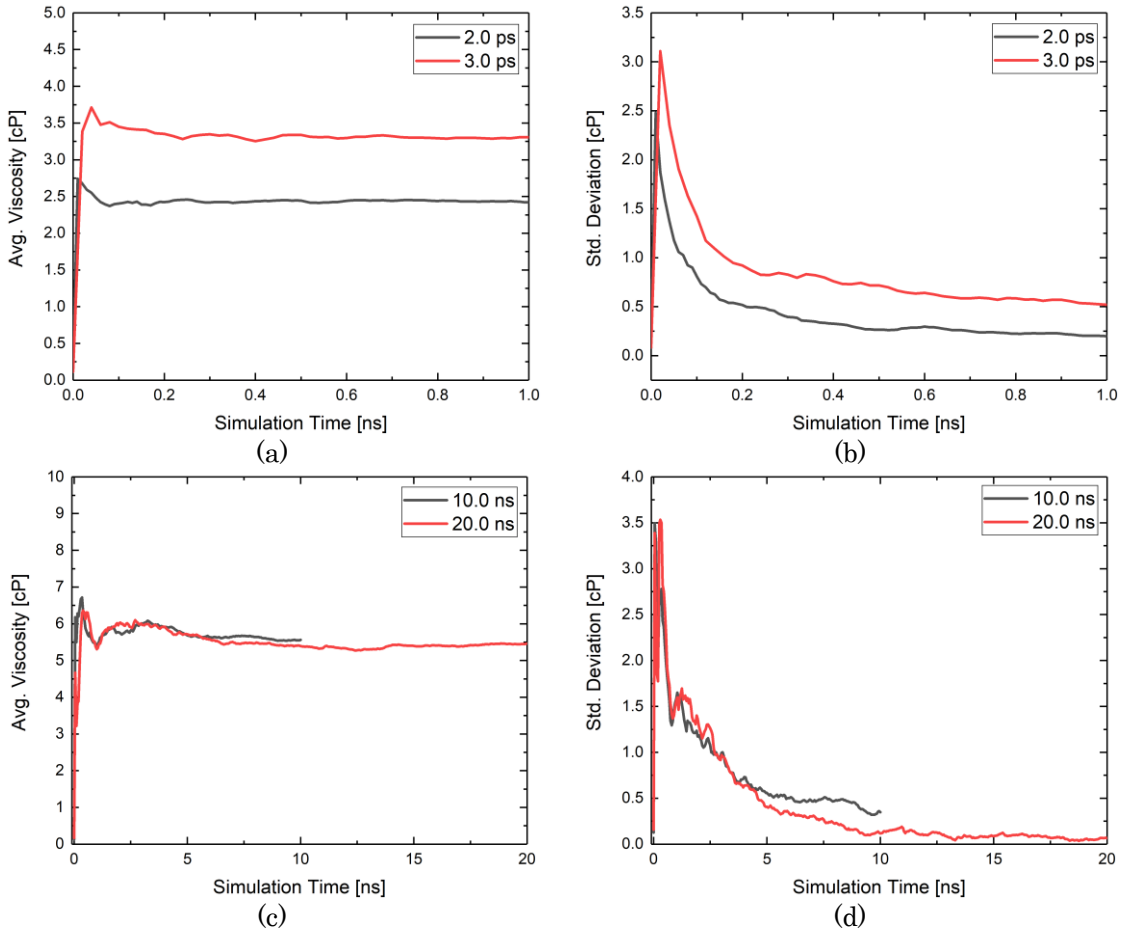


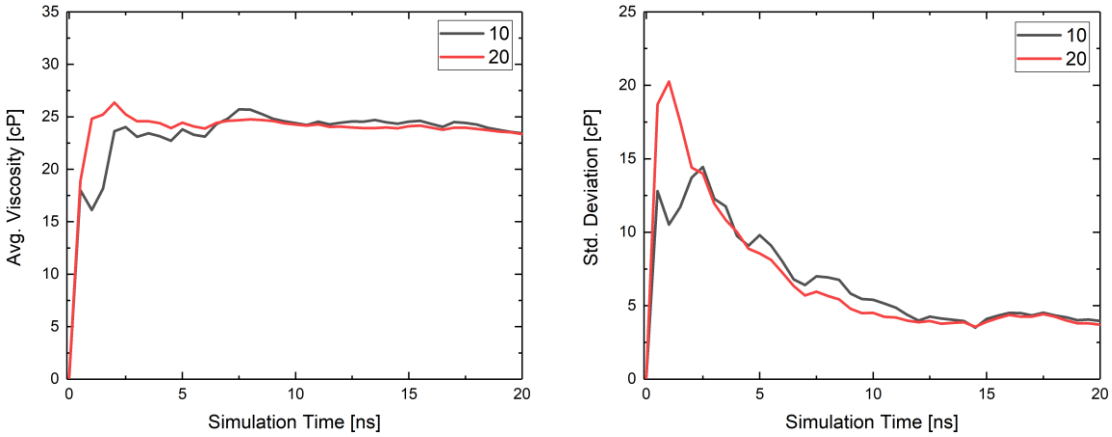
Figure B.5 Optimization of simulation parameters for forcefield at 373.15 K and 1 atm: (1) temperature fluctuation over system size at (a) 40 °C and (b) 100 °C, (2) pressure fluctuation over system size at (c) 40 °C and (d), (e) average temperature for damping coefficients 1, 20,

100, and 200, and (f) average pressure for damping coefficients 1, 20, 100, and 200. This analysis suggested to use around 12000 UA atoms in the simulation box for each fluid to minimize the temperature and pressure fluctuation with damping coefficient of 1 for thermostat and 25 for barostat.

B.7. Dependence of Viscosity on Simulation Parameters

Before simulating viscosity using time decomposition method [160], dependency of viscosity on important simulation parameters, such as correlation length, simulation time, and number of trajectories, were studied to report accurate and reliable viscosity from the simulations [159,160]. Figure B.6 shows the average viscosity and standard deviation as a function of these important parameters. These parameters are arranged here in order of importance from top to bottom.

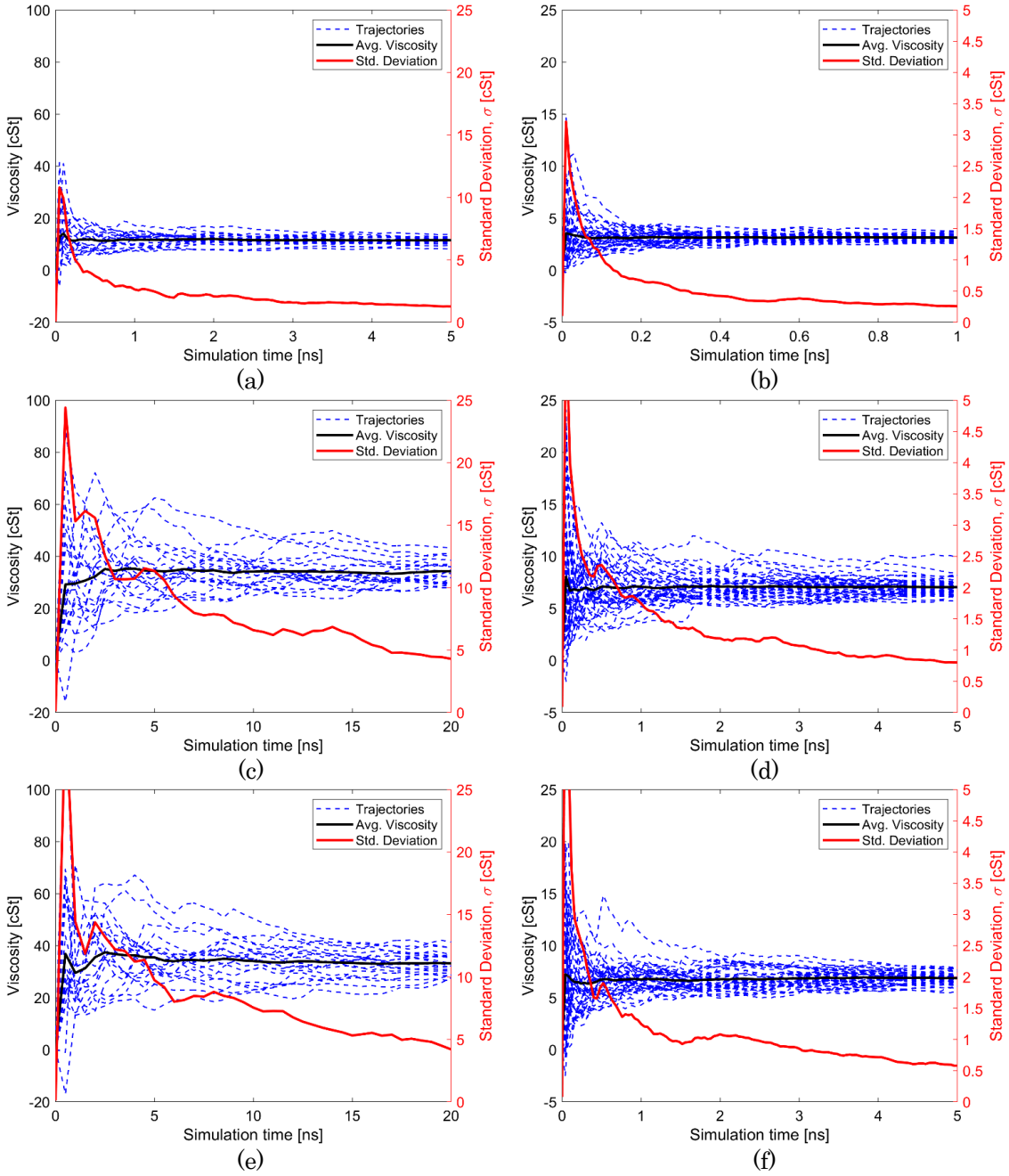




(e) (f)

Figure B.6 Dependence of viscosity on simulation parameters. These results suggest that (1) both average viscosity and standard deviation are highly dependent on the correlation length or time, (2) average viscosity is minutely but std. deviation is highly dependent on the simulation time, and (3) average viscosity and standard deviation are minutely dependent on the number of trajectories, but multiple trajectories are necessary for reporting reliable viscosity.

B.8. Average Simulated Viscosity and Standard Deviation



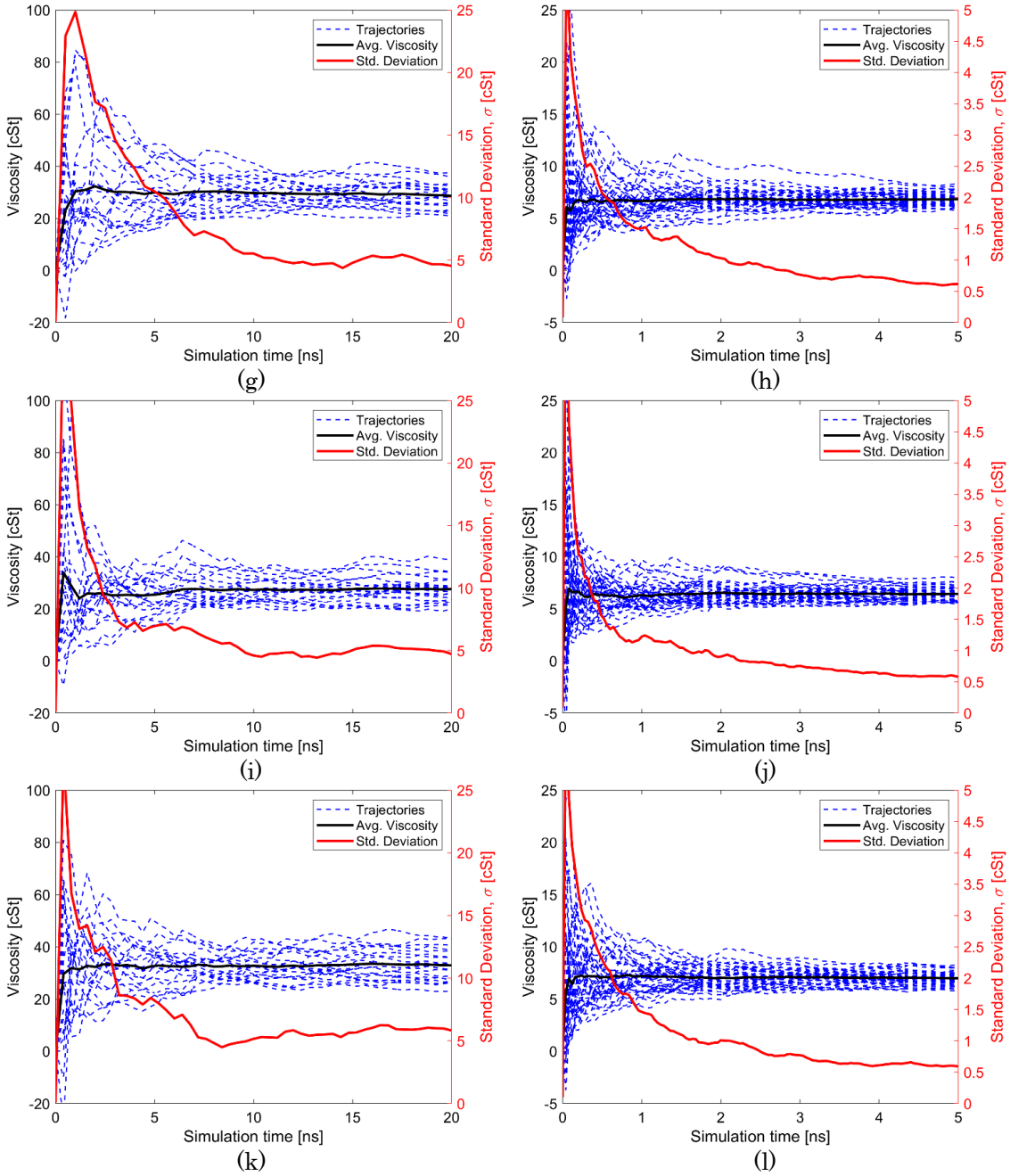
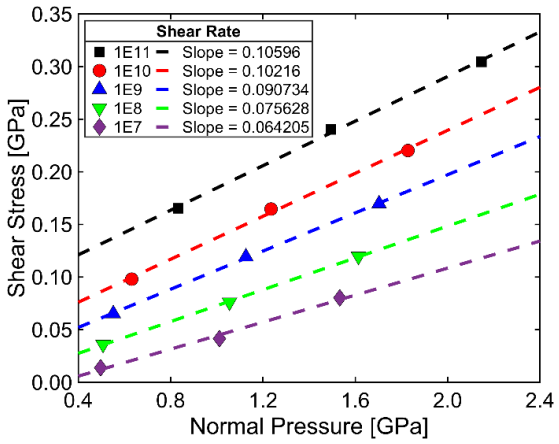
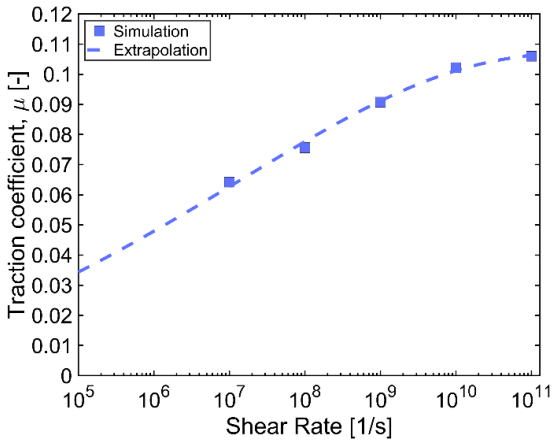


Figure B.7 NVE trajectories with average viscosity and standard deviation curves of the following: (a) PAO 3 cSt at 40 °C, (b) PAO 3 cSt at 100 °C, (c) FKV-PAO at 40 °C, (d) FKV-PAO at 100 °C, (e) FKV-PIB at 40 °C, (f) FKV-PIB at 100 °C, (g) FKV-BDIP at 40 °C, (h) FKV-BDIP at 100 °C, (i) FKV-PAMA at 40 °C, (j) FKV-PAMA at 100 °C, (k) FKV-MABD at 40 °C, and (l) FKV-MABD at 100 °C. Blue dashed lines are NVE trajectories, black and red curves, respectively, represent the average viscosity and standard deviation calculated over these NVE trajectories as a function of simulation time.

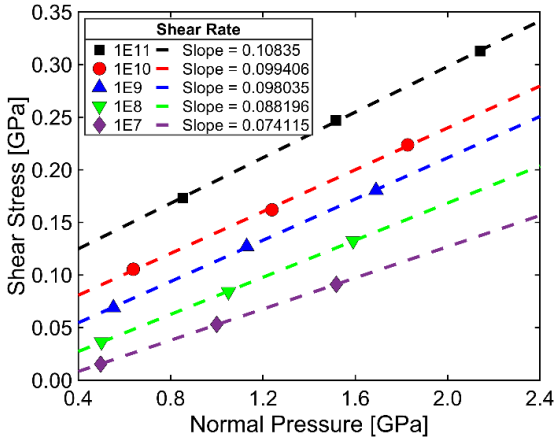
B.9. Traction Coefficient and Extrapolated Traction Curves from Simulations



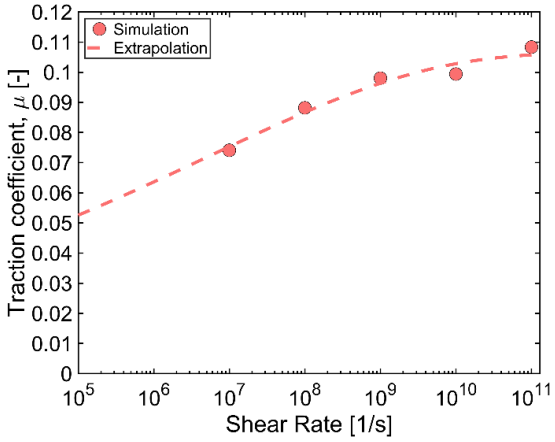
(a)



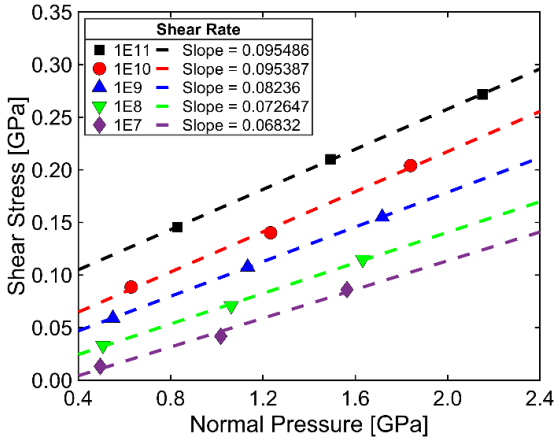
(f)

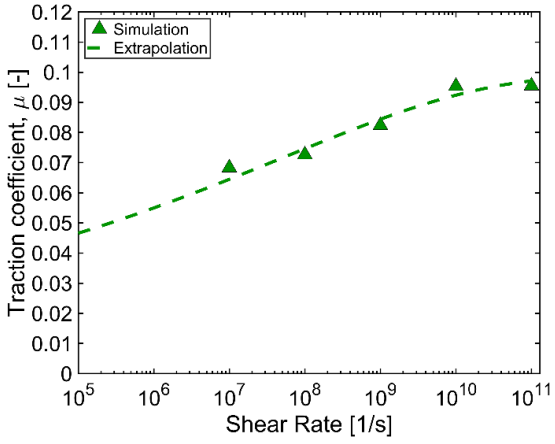


(b)



(g)





(c) (h)

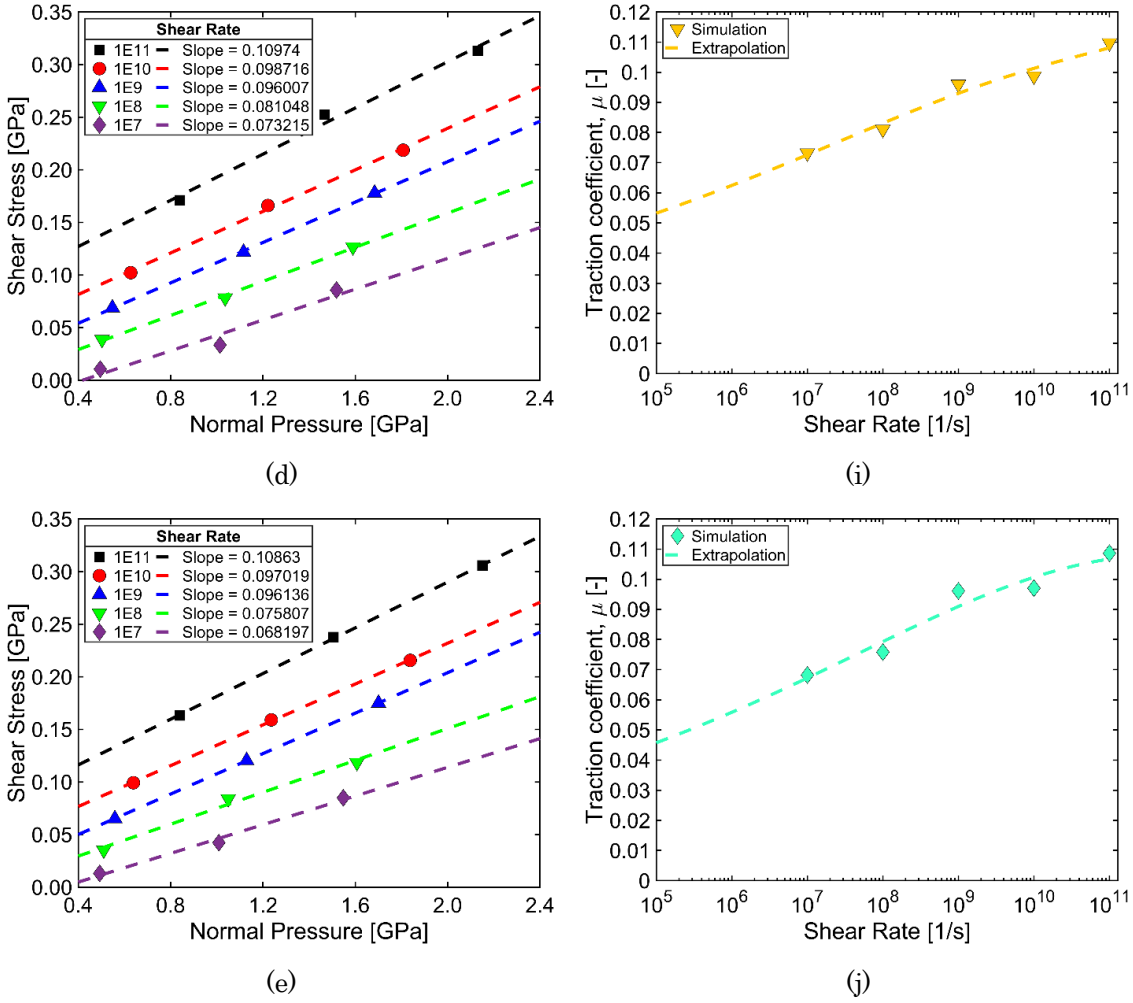


Figure B.8 Traction coefficients at 40 °C directly calculated from the linear fit of the shear stress to the normal load or pressure of the following fluids: (a) FKV-PAO, (b) FKV-PIB, (c) FKV-BDIP, (d) FKV-PAMA, and (e) FKV-MABD. The black, red, blue, and green solid symbols, respectively, represent the shear stress as function of normal pressure at shear rates 1×10^{11}, 1×10^{10}, 1×10^{9}, and 1×10^{8} 1/s. The black, red, blue, and green dash lines, respectively,

represent the linear fit of shear stress to the normal pressure at shear rates 1×10^{11}, 1×10^{10}, 1×10^{9}, and 1×10^{8} 1/s. The slopes of each linear fit are listed with their respective shear rates in the legend box. The extrapolated traction curve at 40 °C of the following fluids: (f) FKV-PAO, (g) FKV-PIB, (h) FKV-BDIP, (i) FKV-PAMA, and (j) FKV-MABD. The solid symbols represent the slopes as the traction coefficients calculated from the linear fit of simulation data of shear stress and normal load. The dashed line curves in these figures represent the extrapolated traction curve obtained by fitting the shear stress and normal pressure to the power function of shear rate.

B.10. Governing Equations

End-to-end distance

In the physical chemistry study of polymers, the end-to-end vector is the vector that points from one end of a polymer to the other end. If each unit in a polymer is represented by a point in space, translation vectors $\vec{r}_i$ connect between these joints. The end-to-end vector $\vec{R}$ is the sum of these translation vectors [189,190]:

$$\vec{R} = \sum_i \vec{r}_i \tag{B.9}$$

The norm of the end-to-end vector is called the end-to-end distance R.

$$R = \sqrt{(x_{1\text{st}} - x_{\text{last}})^2 + (y_{1\text{st}} - y_{\text{last}})^2 + (z_{1\text{st}} - z_{\text{last}})^2} \tag{B.10}$$

Radius of Gyration

In polymer physics, the radius of gyration R_g is used to describe the dimensions of a polymer chain. The radius of gyration of a molecule at a given time is defined as [189,190],

$$R_g = \sqrt{\frac{1}{M}\sum_{i=1}^{n} m_i (r_i - r_{cm})^2} = \sqrt{\frac{1}{M}\sum_{i=1}^{n} m_i [(x_i - x_{cm})^2 + (y_i - y_{cm})^2 + (z_i - z_{cm})^2]} \tag{B.11}$$

Here, $M = \sum_{i=1}^{n} m_i$ and r_{cm}, respectively, are the total mass and center-of-mass position of the molecule and m_i and r_i, respectively, are mass and position of i^{th} atom. The center of mass r_{cm} of the molecule is calculated with the following formula [189,190].

$$r_{cm} = \frac{1}{M}\sum_{i=1}^{n} m_i r_i \tag{B.12}$$

or

$$x_{cm} = \frac{1}{M}\sum_{i=1}^{n} m_i x_i \qquad y_{cm} = \frac{1}{M}\sum_{i=1}^{n} m_i y_i \qquad z_{cm} = \frac{1}{M}\sum_{i=1}^{n} m_i z_i \tag{B.13}$$

Radius of Gyration Tensor

The gyration tensor of the conformations is defined as [189,190],

$$S = \frac{1}{M}\begin{bmatrix} \sum_{i=1}^{n} m_i(x_i - x_{cm})^2 & \sum_{i=1}^{n} m_i(x_i - x_{cm})(y_i - y_{cm}) & \sum_{i=1}^{n} m_i(x_i - x_{cm})(z_i - z_{cm}) \\ \sum_{i=1}^{n} m_i(x_i - x_{cm})(y_i - y_{cm}) & \sum_{i=1}^{n} m_i(y_i - y_{cm})^2 & \sum_{i=1}^{n} m_i(y_i - y_{cm})(z_i - z_{cm}) \\ \sum_{i=1}^{n} m_i(x_i - x_{cm})(z_i - z_{cm}) & \sum_{i=1}^{n} m_i(y_i - y_{cm})(z_i - z_{cm}) & \sum_{i=1}^{n} m_i(z_i - z_{cm})^2 \end{bmatrix} \tag{B.14}$$

Eigenvalues of Gyration Tensor and Shape Parameters

Since the gyration tensor S is a symmetric 3x3 matrix, a Cartesian coordinate system can be found in which it is diagonal (eigen) [189,190],

$$S = \begin{bmatrix} \lambda_x & 0 & 0 \\ 0 & \lambda_y & 0 \\ 0 & 0 & \lambda_z \end{bmatrix} \tag{B.15}$$

Here, the axes are chosen such that the diagonal elements are ordered $\lambda_x \leq \lambda_y \leq \lambda_z$. These diagonal elements are called the principal moments of the gyration tensor [189,190]. The squared radius of gyration is the sum of the principal moments, i.e., $R_g^2 = \lambda_x + \lambda_y + \lambda_z$ [189,190].

The three parameters that describe the shape can determined by the principal moments of the gyration tensor as [189,190],

Acylindricity (c),

$$c = \lambda_z - 0.5(\lambda_y + \lambda_x) \tag{B.16}$$

The acylindricity is always non-negative and zero only when the two principal moments are equal. The condition, $c = 0$ is met when the distribution of particles is cylindrically symmetric which can also be true when the particle distribution is symmetric with respect to the two coordinate axes, e.g., when the particles are distributed uniformly on a regular prism.

Asphericity (b),

$$b = \lambda_y - \lambda_x \tag{B.17}$$

It measures the deviation from the spherical symmetry. The asphericity is always non-negative and zero only when the three principal moments are equal. The condition, $b = 0$ is met when the distribution of particles is spherically symmetric which can also be true when the particles are distributed uniformly on a cube, tetrahedron or other Platonic solid.

Relative shape anisotropy (k),

$$k = \frac{3}{2}\frac{\lambda_x^2 + \lambda_y^2 + \lambda_z^2}{\left(\lambda_x + \lambda_y + \lambda_z\right)^2} - \frac{1}{2} \tag{B.18}$$

It reflects both the symmetry and dimensionality of a polymer conformation. This parameter is limited between the values of 0 (if all points are spherically symmetric) and 1 (if all points lie on a line). It reaches 1 for an ideal linear chain and drops to zero for highly symmetric conformations. For planar symmetric objects, the relative shape anisotropy converges to the value of 1/4 [190].

Shear Rate at Mini Traction Machine (MTM) Test Condition

The regime of lubrication for a point contact was characterized by the dimensionless viscosity g_V and elasticity parameters g_E [11],

$$g_V = \frac{\alpha W^3}{(\eta_0 U)^2 R'^4} = 19.96\text{E}9 \tag{B.19}$$

$$g_E = \frac{W^{8/3}}{(\eta_0 U)^2 E'^{2/3} R'^{10/3}} = 2.27\text{E}8 \tag{B.20}$$

It was assumed that the pressure–viscosity (α) coefficient of the fluids is 15.7 1/GPa [300]. Then, the reduced radius of curvature (R'), reduced young's modulus (E'), contact area dimensions (a), average contact pressures ($\bar{p}$), maximum contact pressure (P_{max}), and maximum shear rate (τ_{max}) were determined using Equations B.21 – B.26 [11], respectively.

$$R' = R_b/2 \tag{B.21}$$

$$E' = 2\left(\frac{1 - v_b^2}{E_b} + \frac{1 - v_d^2}{E_d}\right)^{-1} \tag{B.22}$$

$$a = (3WR'/E')^{1/3} \tag{B.23}$$

$$\bar{p} = \frac{W}{\pi a^2} = 0.86 \text{ GPa} \tag{B.24}$$

$$P_{max} \text{ or } P_H = \frac{3}{2}\bar{p} = 1.29 \text{ GPa} \tag{B.25}$$

$$\tau_{max} = P_{max}/3 = 0.43 \text{ GPa} \tag{B.26}$$

The central film thickness (h_c) for the point contact and compressible liquid can be derived using Hamrock-Dowson formulae as expressed by Equation B.27.

$$h_c = 1.55\alpha^{0.53}(U\eta_0)^{0.67}E'^{0.061}R'^{0.33}P_H^{-0.201} = 0.121 \text{ μm} \tag{B.27}$$

Finally, the shear rate was calculated by Equation B.28.

$$\dot{\gamma} = U\Sigma/h_c = 1.66 \times 10^6 \text{ s}^{-1} \tag{B.28}$$

B.11. Structure Properties of Polymers/Fluids from Molecular Dynamics Simulations

Listed in Table B.10 are the 5 types of molecular weight and 17 structure parameters of the polymers which were studied in this project and, respectively, obtained from the chemistry of the polymers and MD simulations. These parameters were used in modeling VI and TE of the fluids.

Table B.10 All 22 parameters of the polymers at 40 and 100 °C. These were used for VI and TE.

Temperature	40 °C						100 °C					
Parmeter	PAO	PIB	BDIP	PAMA	MABD	ROMP	PAO	PIB	BDIP	PAMA	MABD	ROMP
M_w [kg/mol]	5.10	3.78	8.20	14.00	11.90	10.00	5.10	3.78	8.20	14.00	11.90	10.00
M_n [kg/mol]	5.00	2.10	3.00	8.50	4.30	15.00	5.00	2.10	3.00	8.50	4.30	15.00
M_{wt} [kg/mol]	5.05	2.13	16.26	32.56	11.17	9.88	5.05	2.13	16.26	32.56	11.17	9.88
M_{wb} [kg/mol]	1.01	1.00	13.05	3.34	2.48	5.00	1.01	1.00	13.05	3.34	2.48	5.00
M_{wp} [%]	20.03	47.11	80.24	10.26	22.20	49.57	20.03	47.11	80.24	10.26	22.20	49.57
L_x [Å]	36.29	26.33	109.08	85.43	62.25	107.18	36.05	24.04	94.32	99.20	51.28	97.89
L_y [Å]	40.86	27.35	49.36	52.53	57.33	45.45	35.51	26.31	82.33	56.69	82.70	47.87
L_z [Å]	37.19	23.55	152.31	58.54	67.26	65.65	37.16	24.76	125.89	58.13	75.95	63.36
R [Å]	43.40	29.32	38.70	44.57	68.01	41.14	31.17	26.38	55.92	60.96	87.02	47.63
R_g [Å]	14.78	12.31	51.96	23.96	25.55	35.03	14.13	11.74	43.96	28.18	32.09	33.46
$R_{g,xx}^2$ [Å²]	69.38	56.62	730.20	369.55	213.96	895.32	66.34	42.71	576.33	580.46	119.31	795.94
$R_{g,yy}^2$ [Å²]	83.11	60.04	106.22	100.17	171.97	79.49	60.49	53.39	320.63	113.76	523.50	95.02
$R_{g,zz}^2$ [Å²]	68.35	39.98	1863.34	104.55	268.94	272.60	73.90	46.29	1042.15	100.79	444.16	240.81
$R_{g,xy}^2$ [Å²]	33.75	10.41	-70.70	6.88	54.52	-139.99	14.09	2.42	256.35	-95.18	-42.70	126.52
$R_{g,xz}^2$ [Å²]	13.55	0.98	-851.48	-39.16	-79.30	360.08	-5.14	0.48	-416.27	-35.34	-139.07	220.71
$R_{g,yz}^2$ [Å²]	6.67	11.41	-22.64	-3.20	99.19	-47.06	-26.26	-3.59	-375.48	20.37	302.56	65.22
λ_x [Å²]	160.26	120.76	2320.35	380.53	552.59	1125.32	135.37	105.86	1466.15	608.37	990.78	994.26
λ_y [Å²]	38.88	26.44	301.88	100.69	61.60	76.16	44.31	26.71	337.71	106.03	64.16	87.02
λ_z [Å²]	21.70	9.48	78.41	90.72	40.67	45.92	21.06	9.88	128.82	80.03	32.14	50.49
c [Å²]	129.97	102.86	2130.21	283.66	501.45	1064.27	102.68	87.56	1232.88	515.04	942.63	925.50
b [Å²]	17.19	16.97	223.48	11.18	20.93	30.24	23.24	16.83	208.89	25.34	32.01	36.53

k [-]	0.34	0.41	0.63	0.24	0.59	0.70	0.28	0.37	0.42	0.42	0.69	0.65

Listed in Table B.11 are 17 structure parameters of the polymers at shear rates of 1.6×10^6 1×10^7, 1×10^8, 1×10^9, 1×10^{10}, and 1×10^{11} 1/s. These parameters are the average of three normal pressure 0.5, 1.0, and 1.5 GPa at which TC simulations were executed. The 5 types of molecular weight are the same as listed in Table B.10. These parameters were used in modeling TC of the fluids as a function of shear rate.

Table B.11 All 17 parameters of the polymers at shear rates ranging from 1.6E6 to 1E11 1/s. These were used for TC.

Fluid	FKV-PAO						FKV-PIB					
Shear Rate [1/s]	1E11	1E10	1E9	1E8	1E7	1.6E6	1E11	1E10	1E9	1E8	1E7	1.6E6
L_x [Å]	62.56	53.82	50.49	53.30	48.19	38.03	56.27	49.70	45.83	48.98	45.31	32.61
L_y [Å]	28.33	31.01	34.32	34.94	40.28	32.88	17.18	18.43	19.78	20.36	21.13	23.19
L_z [Å]	19.09	20.21	20.94	22.23	24.47	32.35	14.21	14.75	14.87	15.97	16.48	22.43
R [Å]	45.94	37.54	37.40	44.80	43.65	27.60	46.83	40.43	35.54	42.57	39.28	34.86
R_g [Å]	17.66	15.65	15.28	16.52	16.09	12.86	17.40	15.57	14.77	15.97	15.01	12.81
$R^2_{g,xx}$ [Å²]	270.44	187.36	157.89	188.24	151.65	73.76	298.60	228.01	193.71	232.38	188.49	92.97
$R^2_{g,yy}$ [Å²]	40.53	49.24	65.04	70.08	89.31	47.68	20.37	22.71	28.97	30.27	33.34	36.20
$R^2_{g,zz}$ [Å²]	17.20	18.61	19.24	21.46	26.91	45.80	12.11	13.01	13.08	16.21	17.00	41.90
$R^2_{g,xy}$ [Å²]	0.61	-2.42	1.50	4.41	-9.68	6.95	-1.22	-2.31	1.05	-0.89	-13.17	4.26
$R^2_{g,xz}$ [Å²]	30.56	20.41	15.70	14.45	4.42	19.78	27.45	18.16	15.21	20.21	17.46	3.52
$R^2_{g,yz}$ [Å²]	0.15	-0.23	0.40	-5.15	1.93	0.35	-0.03	0.24	0.47	-0.78	0.96	2.94
λ_x [Å²]	289.24	210.01	192.31	231.74	215.62	103.64	312.19	242.09	212.50	254.52	213.93	142.21
λ_y [Å²]	29.56	34.28	38.10	35.39	37.04	40.90	14.53	16.54	17.75	18.79	18.73	20.82
λ_z [Å²]	9.36	10.91	11.77	12.64	15.21	22.70	4.35	5.10	5.50	5.56	6.17	8.04
c [Å²]	269.78	187.42	167.38	207.73	189.50	71.84	302.75	231.27	200.88	242.34	201.48	127.78
b [Å²]	20.20	23.36	26.33	22.75	21.83	18.20	10.19	11.44	12.25	13.23	12.56	12.79
k [-]	0.62	0.50	0.45	0.53	0.46	0.19	0.74	0.68	0.64	0.65	0.65	0.51
Fluid	FKV-BDIP						FKV-PAMA					
Shear Rate [1/s]	1E11	1E10	1E9	1E8	1E7	1.6E6	1E11	1E10	1E9	1E8	1E7	1.6E6
L_x [Å]	725.72	711.25	927.13	584.46	493.52	141.63	199.95	131.53	127.79	113.41	106.49	86.08
L_y [Å]	45.32	48.12	52.16	52.12	61.68	108.75	42.85	49.10	50.53	51.70	48.66	47.97
L_z [Å]	21.17	21.98	21.52	27.44	39.49	152.42	28.20	35.68	37.64	39.69	47.95	46.40
R [Å]	544.42	552.57	829.83	132.77	179.32	200.85	136.46	76.87	66.84	59.62	55.94	49.06
R_g [Å]	220.10	225.08	295.34	160.33	124.98	57.98	53.48	33.11	32.66	29.73	27.60	24.27
$R^2_{g,xx}$ [Å²]	54848.16	54877.86	88271.42	25619.62	15511.05	1297.90	3000.02	996.17	962.93	738.15	583.59	423.66
$R^2_{g,yy}$ [Å²]	139.97	151.74	168.07	199.76	240.90	516.32	59.04	88.18	100.80	101.27	96.85	94.35
$R^2_{g,zz}$ [Å²]	18.41	17.79	16.40	34.90	74.91	1548.13	30.09	53.03	58.88	61.90	96.20	70.90
$R^2_{g,xy}$ [Å²]	-286.66	711.38	3288.45	727.84	374.06	629.94	14.61	-7.63	68.95	-63.43	-76.30	10.98
$R^2_{g,xz}$ [Å²]	79.54	-31.51	556.94	752.46	877.64	1325.31	100.76	47.65	53.07	-33.47	105.33	-62.40
$R^2_{g,yz}$ [Å²]	2.52	-9.67	22.95	41.81	44.14	722.73	0.60	-0.70	4.49	-0.01	-14.37	-10.74
λ_x [Å²]	54929.80	54954.55	88402.26	25665.22	15577.98	3109.74	3017.42	1017.33	998.48	758.73	629.59	435.19
λ_y [Å²]	64.38	80.65	42.73	179.76	233.63	166.74	50.54	84.77	84.70	98.07	95.54	96.00
λ_z [Å²]	12.37	12.20	10.90	9.30	15.25	85.87	21.20	35.28	39.42	44.52	51.52	57.71
c [Å²]	54891.43	54908.13	88375.45	25570.70	15453.54	2983.43	2981.55	957.31	936.42	687.44	556.06	358.33
b [Å²]	52.01	68.45	31.83	170.47	218.38	80.87	29.35	49.49	45.28	53.55	44.02	38.29
k [-]	0.99	0.99	1.00	0.98	0.95	0.79	0.90	0.67	0.64	0.56	0.48	0.37
Fluid	FKV-MABD						F-ROMP					

Shear Rate [1/s]	1E11	1E10	1E9	1E8	1E7	1.6E6	1E11	1E10	1E9	1E8	1E7	1.9E6
L_x [Å]	145.10	108.06	113.18	91.77	99.36	61.98	313.60	284.76	169.92	218.31	64.56	106.78
L_y [Å]	33.61	40.08	38.67	47.07	45.13	70.46	30.70	34.45	52.17	48.13	64.22	38.47
L_z [Å]	22.97	26.51	26.19	28.85	28.93	41.87	21.71	21.64	26.47	25.64	59.34	44.96
R [Å]	101.04	68.37	68.36	66.60	77.30	59.64	255.14	234.93	110.51	165.40	53.43	38.32
R_g [Å]	41.47	29.83	31.81	25.70	28.07	23.41	92.46	84.49	50.26	62.62	24.16	33.07
$R^2_{g,xx}$ [Å²]	1861.55	905.24	1030.18	570.53	710.20	274.15	9083.97	8518.87	2751.05	5879.82	236.49	964.80
$R^2_{g,yy}$ [Å²]	45.30	71.95	63.74	100.76	93.62	209.71	37.25	62.64	141.08	113.96	178.28	54.41
$R^2_{g,zz}$ [Å²]	22.69	31.46	29.29	38.21	29.77	75.92	21.59	18.92	30.50	29.29	204.41	111.56
$R^2_{g,xy}$ [Å²]	1.23	-15.85	6.80	43.90	-46.05	128.60	46.89	-35.41	17.37	121.04	45.84	46.53
$R^2_{g,xz}$ [Å²]	78.10	40.04	49.02	24.06	-0.25	-29.13	164.20	75.34	73.58	126.62	116.05	162.19
$R^2_{g,yz}$ [Å²]	0.74	-0.88	-5.48	-15.93	-4.99	22.88	-0.12	1.80	-0.93	8.79	85.10	22.02
λ_x [Å²]	1879.98	933.94	1056.14	615.76	742.25	453.49	9101.29	8533.53	2785.67	5916.00	478.61	1029.35
λ_y [Å²]	36.83	57.26	50.15	72.12	69.15	68.48	31.36	54.17	121.44	86.11	95.95	60.48
λ_z [Å²]	12.73	17.45	16.92	21.62	22.20	37.82	10.16	12.74	15.52	20.95	44.63	40.94
c [Å²]	1855.20	896.58	1022.60	568.89	696.58	400.34	9080.54	8500.07	2717.19	5862.47	408.32	978.64
b [Å²]	24.10	39.82	33.23	50.50	46.95	30.66	21.20	41.43	105.92	65.16	51.32	19.55
k [-]	0.86	0.69	0.74	0.56	0.63	0.51	0.98	0.88	0.78	0.74	0.37	0.71

B.12. Test Fluid F-ROMP

To validate developed models of VI, TE, and TC, an addition fluid was formulated "F-ROMP." Its composition and properties are listed in Table B.12. This fluid was formulated by blending 5.8 wt.% of ROMP polymer with the blend of 48.1 wt.% of Nexbase 3.0 cSt and 39.6 wt.% of Nexbase 4.3 cSt lubricants as the base oil. Like the original five fluids, F-ROMP fluids was formulated with 6.5 wt.% of Anglamol 99 additive package. For the simulation purpose, since we did not know the structure of molecules of these Nexbase lubricants, we replaced Nexbase 3.0 cSt with previously formulated PAO 3.0 cSt (see Table B.3) and Nexbase 4.3 cSt with a blend of 84.3 wt.% of PAO 4 cSt and 15.7 wt.% of PAO 8 cSt (viscosity of this blend of PAO 4 cSt and PAO 8 cSt is 4.28 cSt at 100°C = PAO 4.3 cSt). This composition of new base oils results the same viscosity as the blend of Nexbase base oils. Therefore, the F-ROMP fluid model was developed in Material Studio software by mixing 5.8 wt.% of F-ROMP polymer, 51.7 wt.% of PAO3, and 42.5 wt.% of PAO4.3. Figure B.9(a) and B.9(b) shows the simulated and experimental kinematic viscosity both at 40°C and 100°C. These indicate that the new formulated fluid has viscosity close to what we have desired. All the properties of this new test fluid are listed in Table B.13. Figure B.9(c) shows the linear fit of shear stress to the normal pressure of simulation data at 40 °C and shear rates 1×10^{11}, 1×10^{10}, 1×10^{9}, 1×10^{8}, 1×10^{7}, and 1×10^{6} 1/s. Figure B.9(d) shows the extrapolated traction curve at 40 °C of the F-ROMP fluid.

Table B.12 Composition of new test fluid F-ROMP.

Fluid ID		F-ROMP
Polymer	ID	ROMP

$C_{704}H_{1410}$	*Mw* [kg/mol]	15
	PDI	1.5
Polymer [wt.%]		5.8
Base Stock	Nexbase3030 [wt.%] ≈ 3 cSt	48.1
	Nexbase3043 [wt.%] ≈ 4 cSt	39.6
Anglamol 99 additive package [wt.%]		6.5

Table B.13 Measured and simulated properties of new test fluid F-ROMP.

Property	Experimental	Simulation
KV40 [cSt]	25.5	26.1 ± 4.1
KV100 [cSt]	5.6	5.8 ± 0.6
Viscosity Index	168.2	175.3
Thickening Efficiency @ 100°C [%]	9.96	10.35
Density @ 40°C [kg/m³]	-	815.97
Density @ 100°C [kg/m³]	-	778.41

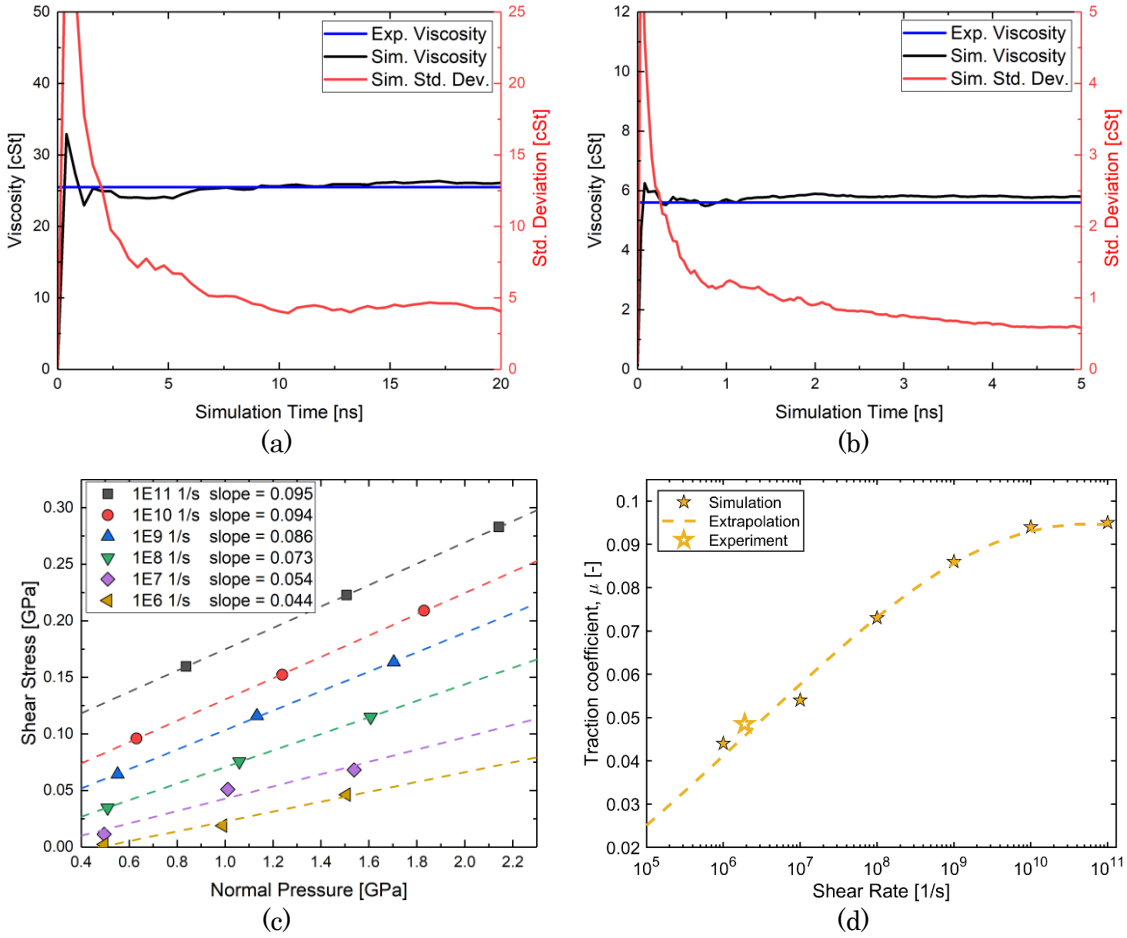


Figure B.9 (a) Kinematic viscosity of fluid F-ROMP at 40 °C, (b) kinematic viscosity of fluid F-ROMP at 100°C, (c) the linear fit of shear stress to the normal pressure of simulation data of F-ROMP, the black, red, blue, green, purple, and orange dash lines, respectively, represent the linear fit of shear stress to the normal pressure at shear rates 1×10^{11}, 1×10^{10}, 1×10^{9}, 1×10^{8}, 1×10^{7}, and 1×10^{6} 1/s. The slopes of each linear fit are listed with their respective shear rates in the legend box, and (d) the extrapolated traction curve of F-ROMP with the simulated and measured traction coefficients shown by the solid and hollow star symbols, respectively. The solid symbols represent the slopes as the traction coefficients calculated from the linear fit of simulation data of shear stress and normal load. The dashed line curves in the figure represents the extrapolated traction curve obtained by fitting the shear stress and normal pressure to the power function of shear rate.

Appendix C.

C.1. Details of Hydrocarbons

Due to the large data set, we used the holdout cross-validation technique. Data were divided randomly into three partitions to develop and assess the models: training, validation, and test data sets. First, 70% of 305 molecules (215 molecules) were used to train the models, and second, 15% of 305 molecules (that is, 45 molecules) were used to validate the models during model development. Third, the remaining 15% of 305 molecules (that is, 45 molecules) that do not belong to the training and validation data sets were selected to assess the accuracy of the estimations with the developed models. Therefore, 70%, 15%, and 15% of 305 molecules were used to train, validate, and test models. All these molecules are listed in Table C.1 with partition IDs 1, 2, and 3 to indicate if that molecule falls in training, validation, and test data sets.

Table C.1 Descriptions of hydrocarbon molecules.

PSU#	Name	Formula	Smiles Code	ID
1	11-n-butyldocosane	$C_{26}H_{54}$	CCCCCCCCCC[C@H](CCCCCCCCCC)CCCC	1
2	9-n-butyldocosane	$C_{26}H_{54}$	CCCCCCCC[C@@H](CCCCCCCCCCCCC)CCCC	1
3	7-n-butyldocosane	$C_{26}H_{54}$	CCCCCC[C@@H](CCCCCCCCCCCCCCC)CCCC	1
4	5-n-butyldocosane	$C_{26}H_{54}$	CCCCC(CCCCCCCCCCCCCCCCC)CCCC	3
5	7-n-hexyldocosane	$C_{28}H_{58}$	CCCCCCC(CCCCCCCCCCCCCCC)CCCCCC	2
6	9-n-octyldocosane	$C_{30}H_{62}$	CCCCCCCCC(CCCCCCCCCCCCC)CCCCCCCC	3
7	11-n-decyldocosane	$C_{32}H_{66}$	CCCCCCCCCCC(CCCCCCCCCCC)CCCCCCCCCC	1
8	11-n-decylheneicosane	$C_{31}H_{64}$	CCCCCCCCCCC(CCCCCCCCCC)CCCCCCCCCC	1
9	11-Phenyl-heneicosane	$C_{27}H_{48}$	c1cccc(c1)C(CCCCCCCCCC)CCCCCCCCCC	1
10	9-p-Tolyloctadecane	$C_{25}H_{44}$	c1(ccc(cc1)[C@@H](CCCCCCCC)CCCCCCCCC)C	3
11	1,1-Dicyclohexyltertradecane	$C_{26}H_{50}$	C1CCC(CC1)C(CCCCCCCCCCCCC)C1CCCCC1	3
12	1,1-Diphenlytetradecane	$C_{26}H_{38}$	c1ccc(cc1)C(c1ccccc1)CCCCCCCCCCCCC	1
13	1,1-Diphenyll-tetradecene	$C_{26}H_{36}$	c1ccc(cc1)C(=CCCCCCCCCCCCC)c1ccccc1	2
15	1-Cyclopentyl-2-hexadecylcyclopentane	$C_{26}H_{50}$	[C@H]1(CCC[C@@H]1CCCCCCCCCCCCCCCC)C1CCCC1	1
16	1-n-Hexadecylindan	$C_{25}H_{42}$	c1cccc2c1CC[C@H]2CCCCCCCCCCCCCCCC	1
18	1-Phenyl-3(2-Pheoylethyl)-hendecane	$C_{25}H_{36}$	c1ccc(cc1)CCC(CCc1ccccc1)CCCCCCCC	3
19	1-Cyclohexyl-3(2-cyclohexylethyl)hendecane	$C_{25}H_{48}$	C1CCC(CC1)CCC(CCCCCCCC)CCC1CCCCC1	1
22	6,11-di-n-amylhexadecane	$C_{26}H_{54}$	CCCCCC(CCCCC(CCCCC)CCCCC)CCCCC	3
23	3-ethyl-5(2-ethylbutyl)octadecane	$C_{26}H_{54}$	C(CCCCCCCCCCCCC)(CC(CC)CC)CC(CC)CC	1
25	9-n-octylheptadecane	$C_{25}H_{52}$	CCCCCCCCC(CCCCCCCC)CCCCCCCC	1
26	9-n-octyl-8-heptadecene	$C_{25}H_{50}$	CCCCCCCC=C(CCCCCCCC)CCCCCCCC	1
27	11-n-amylheneicosane	$C_{26}H_{54}$	CCCCCCCCCCC(CCCCCCCCCC)CCCCC	1
51	7-n-hexyleicosane	$C_{26}H_{54}$	CCCCCCC(CCCCCCCCCCCCC)CCCCCC	1

52	11-Phenyl-10-heneicosene	$C_{27}H_{46}$	c1ccccc1/C(=C\CCCCCCCCC)/CCCCCCCCCC	2
53	11(3-pentyl)heneicosane	$C_{26}H_{54}$	CCCCCCCCCCC(CCCCCCCCCC)C(CC)CC	1
54	11-Benzylheneicosane	$C_{28}H_{50}$	c1ccccc1CC(CCCCCCCCCC)CCCCCCCCCC	1
55	5,14-di-n-butyloctadecane	$C_{26}H_{54}$	CCCCC(CCCCCCCCC(CCCC)CCCC)CCCC	1
58	17-hexadecyltetratriacontane	$C_{50}H_{102}$	CCCCCCCCCCCCCCCCC(CCCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCCCC	1
59	11-20-di-n-decyltriacontane	$C_{50}H_{102}$	CCCCCCCCCCC(CCCCCCCCC(CCCCCCCCCC)CCCCCCCCC)CCCCCCCCCC	3
60	11-Cyclohexylheneicosane	$C_{27}H_{54}$	C1CCCC(C1)C(CCCCCCCCCC)CCCCCCCCCC	1
61	11-alpha-Naphthyl-10-heneicosene	$C_{31}H_{48}$	c1c2c(ccc1)/C(=C/CCCCCCCCC)/CCCCCCCCCC)cccc2	3
62	11-alpha-De-calylheneicosane	$C_{31}H_{60}$	C1[C@H]2[C@@H]([C@@H](CC1)C(CCCCCCCCCC)CCCCCCCCC)CCCC2	1
63	9-n-octyleicosane	$C_{28}H_{58}$	CCCCCCCCC(CCCCCCCCCCC)CCCCCCCC	1
64	11-Cyclopentyl-heneicosane	$C_{26}H_{52}$	C1(CCCC1)C(CCCCCCCCCC)CCCCCCCCCC	1
65	1-Cyclohexyl-2(cyclohexyl-methyl)pentadecane	$C_{28}H_{54}$	C1CCC(CC1)CC(CCCCCCCCCCCCC)CC1CCCCC1	1
67	11-neopentylheneicosane	$C_{26}H_{54}$	CCCCCCCCCCC(CCCCCCCCCC)CC(C)(C)C	1
68	13-Phenylpentacosane	$C_{31}H_{56}$	c1ccccc1C(CCCCCCCCCCCC)CCCCCCCCCCCC	1
69	13-Cyclohexylpentacosane	$C_{31}H_{62}$	C1CCCC(C1)C(CCCCCCCCCCCC)CCCCCCCCCCCC	1
74	11-Cyclopentylmethylheneicosane	$C_{27}H_{54}$	C1(CCCC1)CC(CCCCCCCCCC)CCCCCCCCCC	2
75	3-Cyclohexyleicosane	$C_{26}H_{52}$	C1CCCC(C1)[C@H](CC)CCCCCCCCCCCCCCCCC	1
76	5-Cyclohexyleicosane	$C_{26}H_{52}$	C1CCCCC1[C@@H](CCCC)CCCCCCCCCCCCCCC	3
77	7-Cyclohexyleicosane	$C_{26}H_{52}$	C1CCCCC1[C@@H](CCCCCC)CCCCCCCCCCCCC	2
78	9-Cyclohexyleicosane	$C_{26}H_{52}$	C1CCCCC1[C@@H](CCCCCCCC)CCCCCCCCCC	1
79	3-Phenyleicosane	$C_{26}H_{46}$	c1cccc(c1)[C@@H](CC)CCCCCCCCCCCCCCCCC	1
80	5-Phenyleicosane	$C_{26}H_{46}$	c1cccc(c1)[C@@H](CCCC)CCCCCCCCCCCCCCC	1
81	7-Phenyleicosane	$C_{26}H_{46}$	c1cccc(c1)[C@H](CCCCCCCCCCCCC)CCCCCC	1
82	9-Phenyleicosane	$C_{26}H_{46}$	c1cccc(c1)[C@H](CCCCCCCCCCC)CCCCCCCC	1
87	9(2-Phenylethyl)heptadecane	$C_{25}H_{44}$	c1cccc(c1)CCC(CCCCCCCC)CCCCCCCC	1
88	9(2-Cyclohexylethyl)heptadecane	$C_{25}H_{50}$	C1CCCC(C1)CCC(CCCCCCCC)CCCCCCCC	1
89	1,5-Diphenyl-3(2-pheoylethyl)pentane	$C_{25}H_{28}$	c1ccc(cc1)CCC(CCc1ccccc1)CCc1ccccc1	1
90	1,5-Dicyclohexyl-3(2-cyclopentylpropyl)pentane	$C_{25}H_{46}$	C1CCC(CC1)CCC(CCC1CCCCC1)CCC1CCCCC1	1
91	11-Cyclohexylmethylheneicosane	$C_{28}H_{56}$	C1CCCC(C1)CC(CCCCCCCCCC)CCCCCCCCCC	3
99	1-Phenyleicosane	$C_{26}H_{46}$	c1ccc(cc1)CCCCCCCCCCCCCCCCCCCC	1
100	1-Cyclohexyleicosane	$C_{26}H_{52}$	C1CCC(CC1)CCCCCCCCCCCCCCCCCCCC	1
101	2-Phenyleicosane	$C_{26}H_{46}$	c1ccccc1[C@@H](C)CCCCCCCCCCCCCCCCCC	1
102	2-Cyclohexyleicosane	$C_{26}H_{52}$	C1CCCC(C1)[C@H](CCCCCCCCCCCCCCCCCC)C	1
103	4-Phenyleicosane	$C_{26}H_{46}$	c1ccccc1[C@@H](CCC)CCCCCCCCCCCCCCCC	3
104	4-Cyclohexyleicosane	$C_{26}H_{52}$	C1CCCCC1[C@@H](CCC)CCCCCCCCCCCCCCCC	1
106	n-Hexacosane	$C_{26}H_{54}$	CCCCCCCCCCCCCCCCCCCCCCCCCC	2
107	11-n-decyltetracosane	$C_{34}H_{70}$	CCCCCCCCCCC(CCCCCCCCCCCCC)CCCCCCCCCC	2
108	1-n-Hexadecyloctahydrindan	$C_{25}H_{48}$	C1[C@H]2[C@@H](CCC1)[C@H](CC2)CCCCCCCCCCCCCCC	1

109	3-ethyltetracosane	$C_{26}H_{54}$	CCC(CCCCCCCCCCCCCCCCCCCC)CC	1
110	9(3-Cyclopentylpropyl-heptadecane	$C_{25}H_{50}$	C1(CCCC1)CCCC(CCCCCCCC)CCCCCCCC	1
111	1-Cyclopentyl-4(3-cyclopentylpropyl)dodecane	$C_{25}H_{48}$	C1(CCCC1)CCCC(CCCC1CCCC1)CCCCCCCC	1
112	1,7-Dicyclopentyl-4(3-cyclopentylpropyl)3-heptene	$C_{25}H_{44}$	C1CCC(C1)CCCC(=CCCC1CCCC1)CCCC1CCCC1	3
113	1,7-Dicyclopentyl-4(3-cyclopentylpropyl)heptane	$C_{25}H_{46}$	C(CCCC1CCCC1)(CCCC1CCCC1)CCCC1CCCC1	1
115	1,5-Dicyclopentyl-3(2-cyclopentylethyl) 2-pentene	$C_{25}H_{44}$	C1CCC(CC1)CCC(=CCC1CCCCC1)CCC1CCCCC1	1
116	1,1-Di-p-tolyldo-decane	$C_{26}H_{38}$	c1(ccc(cc1)C(c1ccc(cc1)C)CCCCCCCCCCC)C	1
117	1-Cyclopentyl-henelcosane	$C_{26}H_{52}$	C1(CCCC1)CCCCCCCCCCCCCCCCCCCCC	2
118	2-n-Hexadecylhydrindan	$C_{25}H_{48}$	C1[C@H]2[C@@H](CCC1)C[C@H](C2)CCCCCCCCCCCCCCC	1
119	1,5-Dipthenyl-3(2-Phenylethyl-2-pentene	$C_{25}H_{26}$	c1ccc(cc1)CCC(=CCc1ccccc1)CCc1ccccc1	1
120	2-n-Hexadecylindan	$C_{25}H_{42}$	c1cccc2c1CC(C2)CCCCCCCCCCCCCCCC	1
121	1,10-Di(alphanaphthyl)l-hendecene	$C_{31}H_{34}$	c1cccc2c1cccc2C(=CCCCCCCCCC)c1c2ccccc2ccc1	1
122	1,1-Di(alpha-decalyl)-hendecane	$C_{31}H_{56}$	C1[C@H]2[C@@H]([C@@H](CC1)[C@H]([C@@H]1[C@H]3[C@@H](CCC1)CCCC3)CCCCCCCCCC)CCCC2	1
124	9-n-Dodecyclanthracene	$C_{26}H_{34}$	c1c2c(ccc1)c(c1c(c2)cccc1)CCCCCCCCCCCC	2
125	9-n-Dodecylperhydroanthracene	$C_{26}H_{48}$	C1[C@H]2[C@@H]([C@@H]([C@H]3[C@@H]1CCCC3)CCCCCCCCCCCC)CCCC2	1
126	1,5-Diphenyl-3(3-cyclopentylpropyl)pentane	$C_{25}H_{34}$	c1ccc(cc1)CCC(CCCC1CCCC1)CCc1ccccc1	3
127	1,5-Dicyclohexyl-3(3-cyclopentylpropyl)pentane	$C_{25}H_{46}$	C1CCC(CC1)CCC(CCCC1CCCC1)CCC1CCCCC1	1
128	1,7-Dicyclopentyl-4(2-phenylethyl)heptane	$C_{25}H_{40}$	c1(ccccc1)CCC(CCCC1CCCC1)CCCC1CCCC1	1
129	1,7-Dicyclopentyl-4(2-phenylethyl)heptane	$C_{25}H_{46}$	C1(CCCCC1)CCC(CCCC1CCCC1)CCCC1CCCC1	1
130	1-Phenyl-3(2-cyclohexylethyl)-6-cyclopentylhexane	$C_{25}H_{40}$	c1ccc(cc1)CC[C@H](CCC1CCCCC1)CCCC1CCCC1	1
131	1,10-Di(alphanaphthyl)decane	$C_{30}H_{34}$	c1cccc2c1cccc2CCCCCCCCCCc1c2ccccc2ccc1	2
132	1,10-Di(alphadecalyl)decane	$C_{30}H_{54}$	C1[C@H]2[C@@H]([C@@H](CC1)CCCCCCCCCC[C@@H]1CCC[C@@H]3[C@@H]1CCCC3)CCCC2	1
133	13-n-undecylhexacosane	$C_{36}H_{74}$	CCCCCCCCCCCCCC(CCCCCCCCCCCCC)CCCCCCCCCC	1
134	13-n-dodecylhexacosane	$C_{38}H_{78}$	CCCCCCCCCCCCCC(CCCCCCCCCCCCC)CCCCCCCCCCCC	1
135	15-Phenynonacosane	$C_{35}H_{64}$	c1ccccc1C(CCCCCCCCCCCCCC)CCCCCCCCCCCCCC	1
136	15-Cyclohexylnomacosane	$C_{35}H_{70}$	C1CCCC(C1)C(CCCCCCCCCCCCCC)CCCCCCCCCCCCC	3
137	17-Phenyltritriacontane	$C_{39}H_{72}$	c1ccccc1C(CCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCC	1
138	17-Cyclohexyltritriacontane	$C_{39}H_{78}$	C1CCCC(C1)C(CCCCCCCCCCCCCCCC)CCCCCCCCCCCCCCC	1
139	1,1-Di(4-methylcyclohexyl)dodecane	$C_{26}H_{50}$	[C@H]1(CC[C@@H](CC1)[C@@H](CCCCCCCCCCC)[C@H]1CCC[C@H](C1)C)C	1
140	9-n-dodecylphenanthrene	$C_{26}H_{34}$	c1c2c(ccc1)c1c(c(c2)CCCCCCCCCCCC)cccc1	1
141	9-n-Dodecylperhydrophenanthrene	$C_{26}H_{48}$	C1[C@H]2[C@@H]([C@H]3[C@@H]([C@@H]1CCCCCCCCCCCC)CCCC3)CCCC2	1
142	2-n-Dodecyl-9,10-dihydrophenanthrene	$C_{26}H_{36}$	c1cc2c(-c3c(cc(cc3)CCCCCCCCCCCC)CC2)cc1	1

143	2-n-Dodecylperhydrophenanthrene	$C_{26}H_{48}$	C1[C@H]2[C@@H]([C@H]3[C@@H](C1)C[C@H](CC3)CCCCCCCCCCC)CCCC2	1
144	1,10-Di(5-indanyl)decane	$C_{28}H_{38}$	c12c(ccc(c1)CCCCCCCCCCc1ccc3c(c1)CCC3)CCC2	2
145	1,10-Di(5-hydrindanyl)decane	$C_{28}H_{50}$	[C@@H]12[C@@H](C[C@@H](CC1)CCCCCCCCCC[C@@H]1C[C@H]3[C@@H](CC1)CCC3)CCC2	1
146	2-n-Dodecylphenanthrene	$C_{26}H_{34}$	c1c2c(ccc1)c1c(cc2)cc(cc1)CCCCCCCCCCCC	1
152	1, 4-Di-n-decylbenzene	$C_{26}H_{46}$	c1(ccc(cc1)CCCCCCCCCC)CCCCCCCCCC	1
153	1,4-Di-n-decylcyclohexane	$C_{26}H_{52}$	[C@H]1(CC[C@@H](CC1)CCCCCCCCCC)CCCCCCCCCC	1
155	Cholestane	$C_{27}H_{48}$	C1[C@H]2[C@@](CCC1)([C@H]1[C@@H](CC2)[C@H]2[C@](CC1)([C@@H](CC2)[C@H](C)CCCC(C)C)C)C	2
156	2,4,6-Trimethyln-octadecylbenzene	$C_{27}H_{48}$	c1(cc(c(c(c1)C)CCCCCCCCCCCCCCCCCC)C)C	1
157	1,4,5-Trimethyl-n-octadecylcyclohexane	$C_{27}H_{54}$	[C@H]1(C[C@@H]([C@@H]([C@@H](C1)C)CCCCCCCCCCCCCCCCCC)C)C	2
158	2,5-Dimethyln-octadecylbenzene	$C_{26}H_{46}$	c1cc(c(cc1C)CCCCCCCCCCCCCCCCCC)C	2
159	2,5-Dimethyl-n-octadecylcyclohexane	$C_{26}H_{52}$	C1C[C@@H]([C@@H](C[C@@H]1C)CCCCCCCCCCCCCCCCCC)C	3
161	8-p-Tolylnonadecane	$C_{26}H_{46}$	c1cc(ccc1[C@H](CCCCCCCCCCC)CCCCCCC)C	3
162	8(4-Methylcyclohexyl)nonadecane	$C_{26}H_{52}$	[C@H]1(CC[C@@H](CC1)[C@@H](CCCCCCC)CCCCCCCCCCC)C	1
163	9-n-hexylheptadecane	$C_{23}H_{48}$	CCCCCCCCC(CCCCCCCC)CCCCCC	1
164	9-n-octylhexacosane	$C_{34}H_{70}$	CCCCCCCCC(CCCCCCCCCCCCCCCCC)CCCCCCCC	1
165	1,2,3,4,5,6,7,8,9,10,17,18-Dodecahydro-9(n-octyl)naphthacene	$C_{26}H_{40}$	c1c2c(cc3c1C[C@@H]1[C@H]([C@H]3CCCCCCCC)CCCC1)CCCC2	1
166	9-n-Octylperhydronaphthacene	$C_{26}H_{46}$	C1[C@H]2[C@@H](C[C@H]3[C@H]1CCCC3)[C@H]([C@@H]1[C@H](C2)CCCC1)CCCCCCCC	1
167	11(2,5-Dimethylphenyl)10-heneicosene	$C_{29}H_{50}$	c1cc(c(cc1C)/C(=C/CCCCCCCCC)/CCCCCCCCCC)C	1
168	11(2,5-Dimethylphenyl)heneicosane	$C_{29}H_{52}$	c1cc(c(cc1C)C(CCCCCCCCCC)CCCCCCCCCC)C	2
169	11(2,5-Dimethylcyclohexyl)heneicosane	$C_{29}H_{58}$	[C@H]1(CC[C@@H]([C@@H](C1)C(CCCCCCCCCC)CCCCCCCCC)C)C	2
170	1,7-Diphenyl-4(3-phenylpropyl)3-heptane	$C_{28}H_{32}$	c1ccc(cc1)CCCC(=CCCc1ccccc1)CCCc1ccccc1	2
171	1,7-Diphenyl-4(3-phenylpropyl)heptane	$C_{28}H_{34}$	c1ccc(cc1)CCCC(CCCc1ccccc1)CCCc1ccccc1	1
172	1,7-Dicyclohexyl-4(3-cyclohexylpropyl)heptane	$C_{28}H_{52}$	C1CCC(CC1)CCCC(CCCC1CCCCC1)CCCC1CCCCC1	2
173	11-alpha-ar-Tetralylheneicosane	$C_{31}H_{54}$	c1c2c(c(cc1)C(CCCCCCCCCC)CCCCCCCCCC)CCCC2	1
174	1-alpha-Naphthylpentadecane	$C_{25}H_{38}$	c1c2c(c(cc1)CCCCCCCCCCCCCCC)cccc2	1
175	1-alpha-Decalylpentadecane	$C_{25}H_{48}$	C1[C@H]2[C@@H](CCC1)[C@@H](CCC2)CCCCCCCCCCCCCCC	1
176	n-Octacosane	$C_{28}H_{58}$	CCCCCCCCCCCCCCCCCCCCCCCCCCCC	1
177	9(4-as-perhydroindacenyl)heptadecane	$C_{29}H_{54}$	[C@@H]12[C@H]3[C@@H]([C@@H](C[C@@H]1CCC2)C(CCCCCCCC)CCCCCCCC)CCC3	1
178	9(alpha(cis-0,3,3-Bi-cyclooctyl-methyllheptadecane	$C_{26}H_{50}$	[C@H]1([C@H]2[C@@H](CC1)CCC2)CC(CCCCCCCC)CCCCCCC	1
179	9-n-Octyl(1,2,3,4-tetrahydro)naphthacene	$C_{26}H_{32}$	c1c2c(cc3c1CCCC3)c(c1c(c2)cccc1)CCCCCCCC	1
180	11(2,4-Dimethyleyclopentylmethyl)beneicosane	$C_{29}H_{58}$	[C@H]1(C[C@H](C[C@@H]1C)C)CC(CCCCCCCCCC)CCCCCCCCC	2

181	9(5-Exo-perhydro-4,7-methanoindenylmethyl)-heptadecane	$C_{28}H_{52}$	[C@@H]12[C@@H]3[C@H]([C@@H]([C@@H](C1)CC(CCCCCCC)CC CCCCCC)C2)CCC3	1
182	2,2,4,15,17,17-hexamethyl-7,12-di(3,5,5-trimethylhexyl)octadecane	$C_{42}H_{86}$	[C@H](CCCC[C@@H](CC[C@H](CC(C)(C)C)C)CC[C@H](CC(C)(C)C)C)(CC [C@H](CC(C)(C)C)C)CC[C@@H](CC(C)(C)C)C	1
183	2,2,4,10,12,12-hexamethyl-7(3,5,5-trimethylhexyl)-6-tridecene	$C_{28}H_{56}$	C(=C\C[C@@H](CC(C)(C)C)C)(\CC[C@@H](CC(C)(C)C)C)/CC[C@H](CC(C)(C)C)C	1
184	2,2,4,10,12,12-hexamethyl-7(3,5,5-trimethylhexyl)tridecane	$C_{28}H_{58}$	C(CC[C@H](CC(C)(C)C)C)(CC[C@H](CC(C)(C)C)C)CC[C@H](CC(C)(C)C)C	1
188	1(1-ar-Tetralyl)pentadecane	$C_{25}H_{42}$	c1c2c(c(cc1)CCCCCCCCCCCCCCC)CCCC2	2
189	1(5-Acenaphthyl)pentadecane	$C_{27}H_{40}$	c12c3c(c(cc1)CCCCCCCCCCCCCCC)cccc3CC2	1
190	n-Hexatriacontane	$C_{36}H_{74}$	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1
191	9-n-octyltetracosane	$C_{32}H_{66}$	CCCCCCCCC(CCCCCCCCCCCCCCC)CCCCCCCC	1
192	1-Cyclohexyl-4(alpha-de-calyl)tetradecane	$C_{30}H_{56}$	C1[C@H]2[C@@H](CCC1)[C@@H](CCC2)[C@H](CCCCCCCCC)CCC C1CCCCC1	2
193	1(5-Perhydro-acenaphthyl-pentadecane	$C_{27}H_{50}$	[C@@H]12[C@H]3[C@@H]([C@@H](CC1)CCCCCCCCCCCCCCC)CC C[C@H]3CC2	2
196	6-n-Octylperhydrobenz(de)-anthracene	$C_{25}H_{44}$	C1[C@H]2[C@@H](CCC1)C[C@H]1[C@H](CC[C@@H]3[C@H]1[C@ @H]2CCC3)CCCCCCCC	2
197	n-Dotriacontane	$C_{32}H_{66}$	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1
199	1,3-Dicyclopentyl-2-dodecylcyclopentane	$C_{27}H_{50}$	[C@H]1([C@@H](CC[C@@H]1C1CCCC1)C1CCCC1)CCCCCCCCCCC	1
200	1,1-Di(5-perhydroxcenaphthyl)ethane	$C_{26}H_{42}$	[C@@H]12[C@H]3[C@@H](CCC1)[C@@H](CC[C@@H]3CC2)[C@H] ([C@@H]1[C@H]2CCC[C@@H]3[C@H]2[C@@H](CC1)CC3)C	1
202	1,1-Dicyclopentylhexadecane	$C_{26}H_{50}$	CCCCCCCCCCCCCCCC(C1CCCC1)C1CCCC1	1
203	tri(alphadecalyl)methane	$C_{31}H_{52}$	C1[C@H]2[C@@H](CCC1)[C@@H](CCC2)[C@@H]([C@@H]1CCC[C@ @H]2[C@H]1CCCC2)[C@H]1CCC[C@H]2[C@@H]1CCCC2	1
204	13(alpha-Decalyl)-perhydrodibenzo-(a,i)-fluorene	$C_{31}H_{50}$	[C@@H]12[C@@H]([C@H]3[C@@H]([C@@H]1[C@H]1CC[C@@H] 4[C@@H]1CCCC4)[C@@H]1[C@H](CC3)CCCC1)CC[C@@H]1[C@H]2CCCC1	1
205	n-Tetratetracontane	$C_{44}H_{90}$	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	1
206	1,4-Di(4-phenylbutyl)benzene	$C_{26}H_{30}$	c1ccc(cc1)CCCCc1ccc(cc1)CCCCc1ccccc1	1
207	1,4-Di(4-cyclohexylbutyl)cyclohexane	$C_{26}H_{48}$	[C@@H]1(CC[C@@H](CC1)CCCCC1CCCCC1)CCCCC1CCCCC1	3
208	1,3-Di-n-decylbenzene	$C_{26}H_{46}$	c1cc(cc(c1)CCCCCCCCCC)CCCCCCCCCC	1
209	1,3-Di-n-decylcyclohexane	$C_{26}H_{52}$	[C@H]1(CCC[C@@H](C1)CCCCCCCCCC)CCCCCCCCCC	3
210	9-ethyl-9-n-heptyloctadecane	$C_{27}H_{56}$	CCCCCCCC[C@](CCCCCCCCC)(CC)CCCCCCC	1
211	10-n-heptyl-10-n-octyleicosane	$C_{35}H_{72}$	CCCCCCCCC[C@](CCCCCCCCCC)(CCCCCCC)CCCCCCCC	1
215	3-n-Decylpyrene	$C_{26}H_{30}$	c1cc2c3c(c1CCCCCCCCCC)ccc1c3c(ccc1)cc2	1
216	3-n-Decylperhydropyrene	$C_{26}H_{46}$	[C@H]12[C@H]3[C@H]4[C@H](CC1)CC[C@H]([C@@H]4CC[C@@H] 3CCC2)CCCCCCCCCC	2
218	1-n-Decyl-3,4,5,8,9,10-hexahydropyrene	$C_{26}H_{36}$	c12c3c4c(cc1)CCCc4c(cc3CCC2)CCCCCCCCCC	1

219	4-Decylperhydropyrene*	$C_{26}H_{46}$	[C@H]12[C@H]3[C@H]4[C@H](CC1)CCC[C@@H]4[C@H](C[C@@H]3CCC2)CCCCCCCCCC	1
220	n-Pentatriacontane	$C_{35}H_{72}$	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	3
223	2,6,10,15,19,23-hexamethyltetracosane	$C_{30}H_{62}$	CC(CCC[C@@H](CCC[C@@H](CCCC[C@H](CCC[C@H](CCCC(C)C)C)C)C)C)C	1
224	2-n-Octylchrysene	$C_{26}H_{28}$	c1cc2c(cc1)c1c(c3c(c(c1)CCCCCCCC)cccc3)cc2	3
225	2-Octylperhydrochrysene	$C_{26}H_{46}$	C1[C@H]2[C@@H](CCC1)[C@H]1[C@H](CC2)[C@@H]2[C@@H]([C@@H](C1)CCCCCCCC)CCCC2	2
226	2-Octyltriphenylene*	$C_{26}H_{28}$	c1cc2c(cc1)c1c(c3cc(ccc23)CCCCCCCC)cccc1	3
228	2-Octylperhydrotriphenylene*	$C_{26}H_{46}$	C1[C@H]2[C@@H](CCC1)[C@H]1[C@H]([C@@H]3[C@H]2CC[C@@H](C3)CCCCCCCC)CCCC1	3
229	7-Hexadecylspiro(4,5)decane	$C_{26}H_{50}$	C1CCC2(C[C@@H]1CCCCCCCCCCCCCCCC)CCCC2	1
230	2-Decyl-4b,5,9b,10-tetrahydroindeno(2.1-a)indene*	$C_{26}H_{34}$	c1c2c(ccc1)C[C@H]1[C@@H]2Cc2ccc(cc12)CCCCCCCCCC	2
231	2-Decylperhydroindeno(2.1-a)-indene*	$C_{26}H_{46}$	C1[C@H]2[C@@H](CCC1)C[C@@H]1[C@@H]2C[C@@H]2CC[C@@H](C[C@H]12)CCCCCCCCCC	1
232	11-Octyl-1,2,3,4,5,6,7,8,13,14,15,16-decahydrochrynene*	$C_{26}H_{40}$	c12c(c3c(c(c1)CCCCCCCC)CCCC3)CC[C@@H]1[C@@H]2CCCC1	1
235	1,4-Dimethyl-2-(3,7-Dimethyl-octyl) benzene	$C_{18}H_{30}$	c1cc(c(cc1C)CC[C@@H](CCCC(C)C)C)C	2
236	9-n-Dodecylfluorene*	$C_{25}H_{34}$	c1cccc2c1-c1c(C2CCCCCCCCCCCC)cccc1	1
237	Perhydro-9-n-dodecylfluorene*	$C_{25}H_{46}$	C1[C@@H]2[C@@H](CCC1)[C@H]([C@H]1CCCC[C@@H]21)CCCCCCCCCCCC	3
500	7-n-hexyltridecane	$C_{19}H_{40}$	CCCCCCC(CCCCCC)CCCCCC	1
501	1, 1-Diphenyl-1-heptene	$C_{19}H_{22}$	c1ccc(cc1)C(=CCCCCC)c1ccccc1	1
502	7-Phenyltri-decane	$C_{19}H_{32}$	c1ccc(cc1)C(CCCCCC)CCCCCC	1
503	1, 1-Diphenyl-heptene	$C_{19}H_{24}$	c1ccc(cc1)C(c1ccccc1)CCCCCC	1
504	7-Cyclohexyltridecane	$C_{19}H_{38}$	C1CCC(CC1)C(CCCCCC)CCCCCC	1
505	1,1-Dicyclohexylheptane	$C_{19}H_{36}$	C1CCC(CC1)C(CCCCCC)C1CCCCC1	1
506	7-Phenyl-6-tridecene	$C_{19}H_{30}$	c1ccc(cc1)/C(=C/CCCCC)/CCCCCC	1
507	Tricyclohexylmethane	$C_{19}H_{34}$	C1CCC(CC1)C(C1CCCCC1)C1CCCCC1	1
509	9-Cyclohexyl-heptadecane	$C_{23}H_{46}$	C1CCC(CC1)C(CCCCCCCC)CCCCCCCC	1
510	4-n-propylheptadecane	$C_{20}H_{42}$	CCCC(CCCCCCCCCCCCC)CCC	1
511	5-n-butylhexadecane	$C_{20}H_{42}$	CCCCC(CCCCCCCCCCC)CCCC	1
512	7-methyltridecane	$C_{14}H_{30}$	CCCCCCC(CCCCCC)C	1
513	2-Phenyloctane	$C_{14}H_{22}$	c1cccc(c1)[C@H](CCCCCC)C	2
514	2-cyclohexyloctane	$C_{14}H_{28}$	C1CCC(CC1)[C@@H](C)CCCCCC	1
516	1,1-Diphenyl-ethane	$C_{14}H_{14}$	c1ccc(cc1)C(c1ccccc1)C	3
517	1-Phenyl-1-cyclohexylethane	$C_{14}H_{20}$	c1ccc(cc1)[C@@H](C1CCCCC1)C	1
518	1,1-Dicyclohexylethane	$C_{14}H_{26}$	C1CCC(CC1)C(C)C1CCCCC1	1
519	1,2-Diphenyl-ethane	$C_{14}H_{14}$	c1ccc(cc1)CCc1ccccc1	1
520	1,2-Dicyclohexylethane	$C_{14}H_{26}$	C1CCC(CC1)CCC1CCCCC1	1
521	1-Phenyl-2-cyclohexylethane	$C_{14}H_{20}$	c1ccc(cc1)CCC1CCCCC1	1
522	1-Phenyl-3-cyclopentylpropane	$C_{14}H_{20}$	c1ccc(cc1)CCCC1CCCC1	1

523	1-Cyclohexyl-3-cyclopentylpropane	$C_{14}H_{26}$	C1CCC(CC1)CCCC1CCCC1	1
524	diphenylmethane	$C_{13}H_{12}$	c1ccc(cc1)Cc1ccccc1	1
525	Dicyclohexylmethane	$C_{13}H_{24}$	C1CCC(CC1)CC1CCCCC1	3
526	9-n-Butylanthracene	$C_{18}H_{18}$	c1cc2c(cc1)c(c1c(c2)cccc1)CCCC	2
527	9-n-Butylperhydroanthracene	$C_{18}H_{32}$	C1[C@H]2[C@@H](CCC1)[C@H]([C@@H]1[C@H](C2)CCCC1)CCCC	3
528	n-Dodecane	$C_{12}H_{26}$	CCCCCCCCCCCC	1
529	n-Tridecane	$C_{13}H_{28}$	CCCCCCCCCCCCC	1
530	1-Tridecene	$C_{13}H_{26}$	C=CCCCCCCCCCCC	1
531	n-Tetradecane	$C_{14}H_{30}$	CCCCCCCCCCCCCC	1
532	n-Pentadecane	$C_{15}H_{32}$	CCCCCCCCCCCCCCC	1
533	1-Pentadecene	$C_{15}H_{30}$	C=CCCCCCCCCCCCCC	1
534	n-Hexadecane	$C_{16}H_{34}$	CCCCCCCCCCCCCCCC	1
535	n-Heptadecane	$C_{17}H_{36}$	CCCCCCCCCCCCCCCCC	1
536	1-Heptadecene	$C_{17}H_{34}$	C=CCCCCCCCCCCCCCCC	1
537	n-Octadecane	$C_{18}H_{38}$	CCCCCCCCCCCCCCCCCC	3
538	1-Phenyloctane	$C_{14}H_{22}$	c1ccc(cc1)CCCCCCCC	1
539	1-cyclohexyloctane	$C_{14}H_{28}$	C1CCC(CC1)CCCCCCCC	1
540	n-Eicosane	$C_{20}H_{42}$	CCCCCCCCCCCCCCCCCCCC	1
541	n-Tetracosane	$C_{24}H_{50}$	CCCCCCCCCCCCCCCCCCCCCCCC	1
542	7-Cyclopentylmethyltridecane	$C_{19}H_{38}$	C1CCC(C1)CC(CCCCCC)CCCCCC	1
543	cis(0,3,3)Bicyclo-octane	C_8H_{14}	C1[C@H]2[C@@H](CC1)CCC2	1
544	1-alpha-Decalylhendecane	$C_{21}H_{40}$	C1[C@H]2[C@@H]([C@@H](CC1)CCCCCCCCCCC)CCCC2	1
545	7-n-propyltridecane	$C_{16}H_{34}$	CCCCCCC(CCCCCC)CCC	3
546	5-n-butylnonane	$C_{13}H_{28}$	CCCCC(CCCC)CCCC	3
547	5-n-butyl-4-nonene	$C_{13}H_{26}$	CCC/C(=C/CCCC)/CCCC	1
548	1,3-Dicyclopentylcyclopentane	$C_{15}H_{26}$	C1CCC(C1)[C@H]1CC[C@@H](C1)C1CCCC1	1
549	4-n-propylheptane	$C_{10}H_{22}$	CCCC(CCC)CCC	1
550	4-n-propyl-3-heptene	$C_{10}H_{20}$	CCC=C(CCC)CCC	3
551	Bicyclopentyl	$C_{10}H_{18}$	C1C(CCC1)C1CCCC1	1
552	1,5-Dicyclopentyl-3(2-cyclopentylethyl) 2-pentene	$C_{22}H_{38}$	C1CC(CC1)CCC(=CCC1CCCC1)CCC1CCCC1	2
553	1,5-Dicyclopentyl-3(2-cyclopentylethyl) 2-pentene	$C_{22}H_{40}$	C(CCC1CCCC1)(CCC1CCCC1)CCC1CCCC1	1
554	8-n-hexylpentadecane	$C_{21}H_{44}$	CCCCCCCC(CCCCCCC)CCCCCC	1
555	2,2,3,5,5,6,6-heptamethyl-3-heptene	$C_{14}H_{28}$	CC(/C(=C/C(C(C)(C)C)(C)C)(C)C)/C)(C)C	3
556	2,2,3,3,5,6,6-heptamethylheptane	$C_{14}H_{30}$	CC([C@H](CC(C(C)(C)C)(C)C)(C)C)C)(C)C	3
557	4-9-di-n-propyldodecane	$C_{18}H_{38}$	CCCC(CCCCC(CCC)CCC)CCC	3
559	1-alpha-Napthylbendecane	$C_{21}H_{30}$	c1c2c(ccc1)c(ccc2)CCCCCCCCCCC	1
560	1,1,-diphenylethylene	$C_{14}H_{12}$	c1ccc(cc1)/C=C/c1ccccc1	1
561	Perhydrofluorene	$C_{13}H_{22}$	C1[C@@H]2[C@@H](CCC1)C[C@H]1CCCC[C@@H]21	1

562	1,2-Di(alphadecalyl)ethane	$C_{22}H_{38}$	C1[C@@H]2[C@@H]([C@@H](CC1)CC[C@@H]1[C@H]3[C@@H](CCC1)CCCC3)CCCC2	3
563	1,1-Di(alphadecalyl)ethane	$C_{22}H_{38}$	C1[C@H]2[C@@H](CCC1)[C@@H](CCC2)[C@@H](C)[C@H]1[C@@H]2[C@@H](CCCC2)CCC1	2
564	Tricyclopentylmethane	$C_{16}H_{28}$	C1CCC(C1)C(C1CCCC1)C1CCCC1	2
565	1-Undecene	$C_{11}H_{22}$	C=CCCCCCCCCC	1
566	1-Dodecene	$C_{12}H_{24}$	C=CCCCCCCCCCC	1
567	1-Methylnaphthalene	$C_{11}H_{10}$	c1cc2c(cc1)c(ccc2)C	2
568	2-Methylnaphthalene	$C_{11}H_{10}$	c1cc2c(cc1)cc(cc2)C	1
569	cis-Decahydro-naphthalene	$C_{10}H_{18}$	C1[C@@H]2[C@@H](CCC1)CCCC2	1
570	trans-Decahydro-naphthalene	$C_{10}H_{18}$	C1[C@@H]2[C@@H](CCC1)CCCC2	1
571	1-Phenyldecane	$C_{16}H_{26}$	c1ccc(cc1)CCCCCCCCCC	2
572	1-Cyclohexyldecane	$C_{16}H_{32}$	C1CCC(CC1)CCCCCCCCCC	1
573	1-Cyclopentyldecane	$C_{15}H_{30}$	C1CCC(C1)CCCCCCCCCC	3
574	1 2,3,4,5,6,7,8,13,14,15,16-Dodecahydrochrysene	$C_{18}H_{24}$	c12c(c3c(cc1)CCCC3)CC[C@@H]1[C@H]2CCCC1	1
575	Perhydrochrysene	$C_{18}H_{30}$	C1[C@H]2[C@@H]([C@H]3[C@@H](C1)CC[C@@H]1[C@@H]3CCCC1)CCCC2	1
576	1,2,3,4-Tetrahydrofluoranthene	$C_{16}H_{14}$	c1cc2c3c(-c4c([C@@H]3CCC2)cccc4)c1	1
577	Perhydrofluoranthene	$C_{16}H_{26}$	C1[C@H]2[C@H]3[C@@H](CC1)[C@@H]1[C@H]([C@@H]3CCC2)CCCC1	1
578	Perhydropyrene	$C_{16}H_{26}$	C1C[C@H]2[C@H]3[C@@H](C1)CC[C@@H]1[C@@H]3[C@@H](CCC1)CC2	1
580	1,1-Dicyclopentylethane	$C_{12}H_{22}$	C1CC(CC1)C(C)C1CCCC1	2
581	2-methyldecane	$C_{11}H_{24}$	CC(CCCCCCCC)C	1
582	2-methylpentadecane	$C_{16}H_{34}$	CC(CCCCCCCCCCCCC)C	3
583	2-methylheptadecane	$C_{18}H_{38}$	CC(CCCCCCCCCCCCCCC)C	1
584	2-methyltricosane	$C_{24}H_{50}$	CC(CCCCCCCCCCCCCCCCCCCCC)C	2
586	Di(alpha -decalyl)-methane	$C_{21}H_{36}$	C1[C@H]2[C@@H](CCC1)[C@@H](CCC2)C[C@@H]1CCC[C@H]2[C@@H]1CCCC2	2
587	Perhydrodibenzo-(a,i)fluorene	$C_{21}H_{34}$	C1[C@@H]2[C@@H]([C@H]3[C@@H](C1)CCCC3)C[C@H]1[C@@H]3[C@H](CC[C@@H]21)CCCC3	3
588	3-methyleicosane	$C_{21}H_{44}$	CC[C@H](CCCCCCCCCCCCCCCCC)C	2
589	1-Tetradecene	$C_{14}H_{28}$	C=CCCCCCCCCCCC	1
590	1-Hexadecene	$C_{16}H_{32}$	C=CCCCCCCCCCCCCC	2
591	10-methyleicosane	$C_{21}H_{44}$	CCCCCCCCC[C@@H](CCCCCCCCCC)C	1
592	2(ar)-Butyltetralin	$C_{14}H_{20}$	c1cc2c(cc1CCCC)CCCC2	1
593	2(ar)-Decylitetralin	$C_{20}H_{32}$	c1c2c(cc(c1)CCCCCCCCCC)CCCC2	2
594	2-Decyldacalin	$C_{20}H_{38}$	C1[C@H]2[C@@H](C[C@@H](C1)CCCCCCCCCC)CCCC2	1
595	2-Decylindan	$C_{19}H_{30}$	c1cc2c(cc1)CC(C2)CCCCCCCCCC	1
596	2-Decylhydrindan	$C_{19}H_{36}$	C1[C@@H]2[C@@H](CCC1)C[C@H](C2)CCCCCCCCCC	1
597	5-Decylindan	$C_{19}H_{30}$	c1cc2c(cc1CCCCCCCCCC)CCC2	1
598	5-Decylhydrindan	$C_{19}H_{36}$	C1[C@@H]2[C@@H](C[C@@H](C1)CCCCCCCCCC)CCC2	1

599	2-Butyl-lhexylindene	$C_{19}H_{28}$	c1cc2c(cc1)C(=C(C2)CCCC)CCCCCC	1
600	2-Butyl-1-hexylindan	$C_{19}H_{30}$	c1cc2c(cc1)[C@@H]([C@H](C2)CCCC)CCCCCC	1
601	2-Butyl-l-hexylhydrindan	$C_{19}H_{36}$	C1[C@H]2[C@@H](CCC1)[C@H]([C@H](C2)CCCC)CCCCCC	3
602	2-Butyl-5-hexylindan	$C_{19}H_{30}$	c1cc2c(cc1CCCCCC)C[C@H](C2)CCCC	1
603	2-Butyl-5-hexylhydrindan	$C_{19}H_{36}$	C1[C@H]2[C@@H](C[C@@H](C1)CCCCCC)C[C@@H](C2)CCCC	1
604	5-Butyl-6-hexylindan	$C_{19}H_{30}$	c1(cc2c(cc1CCCC)CCC2)CCCCCC	2
605	5-Butyl-6-hexylindan	$C_{19}H_{36}$	C1[C@H]2[C@@H](C[C@@H]([C@@H]1CCCCCC)CCCC)CCC2	3
606	2-n-Butylnaphthalene	$C_{14}H_{16}$	c1cc2c(cc1CCCC)cccc2	3
607	2-n-Butyldecalin	$C_{14}H_{26}$	C1[C@H]2[C@@H](CCC1)C[C@H](CC2)CCCC	1
608	Bicyclohexyl	$C_{12}H_{22}$	C1C(CCCC1)C1CCCCC1	3
609	n-Nonadecane	$C_{19}H_{40}$	CCCCCCCCCCCCCCCCCCC	1
610	7-n-Butyl-l-n-hexyloaphthalene	$C_{20}H_{28}$	c1c2c(c(cc1)CCCCCC)cc(cc2)CCCC	1
611	7(ar)-n-Butyl-1-n-hexyltetralin	$C_{20}H_{32}$	c1c2c(cc(c1)CCCC)[C@H](CCC2)CCCCCC	3
612	7-n-Butyl-l-n-hexydecalin	$C_{20}H_{38}$	C1[C@H]2[C@@H]([C@@H](CC1)CCCCCC)C[C@H](CC2)CCCC	3
613	2-n-Butyl-3-n-hexyloaphthalene	$C_{20}H_{28}$	c1c2c(ccc1)cc(c(c2)CCCCCC)CCCC	1
614	2(ar)-n-Butyl-3(ar)-n-hexyltetralin	$C_{20}H_{32}$	c1c2c(cc(c1CCCC)CCCCCC)CCCC2	1
615	2-n-Butyl-3-n-hexyldecalin	$C_{20}H_{38}$	C1[C@H]2[C@@H](CCC1)C[C@H]([C@H](C2)CCCCCC)CCCC	1
616	4,5-Dimethylphenanthrene	$C_{16}H_{14}$	c1cc2c(c(c1)C)c1c(cc2)cccc1C	3
617	4,5-Dimethyl-9,10-dihydrophenanthrene	$C_{16}H_{16}$	c1cc2c(-c3c(CC2)cccc3C)c(c1)C	1
618	4,5-Dimethylperhydrophenanthrene	$C_{16}H_{28}$	C1C[C@H]2[C@@H]([C@@H](C1)C)[C@H]1[C@@H](CCC[C@@H]1C)CC2	2
619	n-Tricosane	$C_{23}H_{48}$	CCCCCCCCCCCCCCCCCCCCCCC	1
620	Spiro(4,5)decane	$C_{10}H_{18}$	C12(CCCC1)CCCCC2	1
622	Spiro(5,5)undecane	$C_{10}H_{20}$	C1CCC2(CC1)CCCCC2	1
623	Spiro(5,6)dodecane	$C_{12}H_{22}$	C1CCC2(CC1)CCCCCC2	1
625	1,2,3,4,5,6,7,8-Octahydrophenanthrene	$C_{14}H_{18}$	c1cc2c(c3c1CCCC3)CCCC2	2
626	Perhydrophenanthren	$C_{14}H_{24}$	C1[C@H]2[C@@H](CCC1)[C@H]1[C@H](CC2)CCCC1	1
627	2,6-Dimethy-3-octylnapthalene	$C_{20}H_{28}$	c1c2c(ccc1C)cc(c(c2)CCCCCCCC)C	1
628	2(ar),6-Dimethyl-3-octyltetralin	$C_{20}H_{32}$	c1c2c(cc(c1C)CCCCCCCC)C[C@H](CC2)C	1
629	2,6-Dimethyl-3-octyldecalin	$C_{20}H_{38}$	C1[C@H]2[C@@H](C[C@@H]([C@@H]1C)CCCCCCCC)C[C@H](CC2)C	1
631	1,3-Diphenyl-benzene	$C_{18}H_{14}$	c1ccc(cc1)c1cccc(c1)c1ccccc1	2
632	1,3-Dicyclopentylcyclopentane	$C_{18}H_{32}$	C1C[C@@H](C[C@@H](C1)C1CCCCC1)C1CCCCC1	1
633	1,2-Diphenyl-benzene	$C_{18}H_{14}$	c1ccc(cc1)c1ccccc1c1ccccc1	1
634	1,2-Dicyclohexylcyclohexane	$C_{18}H_{32}$	C1C[C@@H]([C@@H](CC1)C1CCCCC1)C1CCCCC1	1
637	Perhydroanthracene	$C_{14}H_{24}$	C1[C@H]2[C@@H](CCC1)C[C@H]1CCCC[C@H]1C2	1
638	1,4-Dimethyl-5-octylnaphthalene	$C_{20}H_{28}$	c1(c2c(c(cc1)C)c(ccc2)CCCCCCCC)C	1
639	1(ar),4-Dimethyl-5-octyltetralin	$C_{20}H_{32}$	c1(c2c(c(cc1)C)[C@H](CCC2)CCCCCCCC)C	1
640	1,4-Dimethyl-5-octyldecalin	$C_{20}H_{38}$	C1[C@H]2[C@@H]([C@@H](CC1)CCCCCCCC)[C@@H](CC[C@H]2C)C	1

641	1,2,2a,3,4,5-Hexahydropyrene	$C_{16}H_{16}$	c12c3c4c(cc1)cccc4CC[C@@H]3CCC2	2
642	2,6,10,14-tetramethylpentadecane	$C_{19}H_{40}$	CC(CCC[C@@H](CCC[C@@H](CCCC(C)C)C)C)C	1
643	2,6,11,15-tetramethylhexadecane	$C_{20}H_{42}$	CC(CCC[C@@H](CCCC[C@H](CCCC(C)C)C)C)C	1
645	Perhydroindeno(2,1-a)indene*	$C_{16}H_{26}$	C1C[C@@H]2[C@@H](CC1)C[C@H]1[C@H]2C[C@H]2CCCC[C@H]12	1
646	1-t-Butylnaphthalene	$C_{14}H_{16}$	c1cc2c(cc1)c(ccc2)C(C)(C)C	1
647	1-t-Butyldecahydronaphthalene*	$C_{14}H_{26}$	C1[C@H]2[C@@H](CCC1)[C@@H](CCC2)C(C)(C)C	3
648	1,2,3,4,4a,9,10,10a-Octahydro-3-hexly-6-butylphenanthrene*	$C_{24}H_{36}$	c1c2c(cc(c1)CCCC)[C@H]1[C@H](CC2)CC[C@@H](C1)CCCCCC	1
652	2-t-Butylnaphthalene	$C_{14}H_{16}$	c1cc2c(cc1)cc(cc2)C(C)(C)C	1
654	Perhydro-1,2,5,6-dibenzcyclo-octane*	$C_{16}H_{28}$	C1[C@H]2[C@@H](CCC1)CC[C@H]1CCCC[C@H]1CC2	1

C.2. Experimental Data

The experimental dynamic viscosity and density of all the 305 hydrocarbons are listed in Table C.2 from 0 °C to 135 °C. For these molecules, the dynamic viscosity and density range from 0.29 cP to 2.00×10^4 cP and 0.67 g/cc to 1.12 g/cc, respectively.

Table C.2 Experimental dynamic viscosity and density of the hydrocarbon molecules from API Project 42 [263].

PSU #	Dynamic Viscosity [cP]					Density [g/cc]				
	32	68	100	140	210	32	68	100	140	210
1	37.46	15.24	8.2	4.466	2.0930	0.8172	0.8041	0.7924	0.7780	0.7529
2	-	15.6	8.389	4.56	2.1200	0.8178	0.8043	0.7930	0.7785	0.7530
3	41.03	16.44	8.774	4.74	2.2000	0.8183	0.8042	0.7928	0.7784	0.7532
4	-	17.11	9.09	4.94	2.2800	0.8194	0.8059	0.7946	0.7796	0.7548
5	-	20.11	10.41	5.59	2.5100	0.8210	0.8080	0.7969	0.7820	0.7571
6	-	22.19	11.5	6.05	2.7000	0.8246	0.8114	0.7999	0.7855	0.7609
7	67.96	25.81	13.21	6.86	2.9800	0.8263	0.8129	0.8017	0.7876	0.7630
8	-	23.53	12.06	6.33	2.7900	0.8245	0.8117	0.8005	0.7862	0.7616
9	-	26.34	12.84	6.42	2.75	0.8661	0.8533	0.8417	0.8272	0.8019
10	78.08	26.15	12.4	6.04	2.52	0.8683	0.8551	0.8435	0.8285	0.8026
11	-	-	31.3	12.6	4.37	0.8858	0.8734	0.8622	0.8481	0.8241
12	-	37.83	16.81	7.89	3.21	0.9321	0.9187	0.9069	0.8923	0.8665
13	-	38.48	17.63	8.08	3.23	0.9407	0.9268	0.9144	0.8987	0.8726
15	-	24.8	12.91	6.86	3.08	0.8720	0.8596	0.8479	0.8337	0.8092
16	-	-	13.21	6.78	3.02	0.8985	0.8854	0.8739	0.8589	0.8335
18	75.77	24.13	11.44	5.76	2.49	0.9343	0.9212	0.9094	0.8949	0.8690
19	291.4	72.84	28.57	11.68	4.05	0.8789	0.8659	0.8548	0.8405	0.8158
22	53.62	19.13	9.358	4.8	2.0600	0.8210	0.8070	0.7954	0.7801	0.7552
23	52.07	18.7	9.444	4.9	2.1900	0.8251	0.8111	0.8000	0.7852	0.7597

25	30.81	12.81	7.04	3.9	1.8600	0.8152	0.8019	0.7905	0.7761	0.7507
26	24.68	10.65	6.07	3.51	1.7300	0.8225	0.8086	0.7965	0.7818	0.7566
27	36.7	14.86	7.963	4.37	2.0500	0.8176	0.8040	0.7924	0.7778	0.7528
51	-	15.73	8.447	4.57	2.1400	0.8172	0.8042	0.7925	0.7782	0.7531
52	61	20.75	10.63	5.54	2.47	0.8772	0.8638	0.8522	0.8378	0.8115
53	39.22	15.53	8.267	4.47	2.0800	0.8225	0.8096	0.7981	0.7837	0.7588
54	58.38	21.58	10.94	5.688	2.535	0.8726	0.8594	0.8474	0.8330	0.8074
55	-	19.72	9.678	4.92	2.14	0.8208	0.8077	0.7970	0.7816	0.7558
58	-	-	40.9	18.93	7.1350	0.8389	0.8268	0.8160	0.8024	0.7788
59	-	-	43.96	19.6	7.0130	0.8394	0.8273	0.8165	0.8029	0.7792
60	89.33	30.14	14.44	7.07	2.94	0.8501	0.8375	0.8261	0.8121	0.7875
61	282.2	76.36	31.22	13.22	4.61	0.9167	0.9037	0.8923	0.8780	0.8523
62	403.7	101.5	39.09	15.53	5.2	0.8824	0.8699	0.8585	0.8445	0.8203
63	-	18.03	9.548	5.09	2.3400	0.8205	0.8076	0.7961	0.7815	0.7567
64	53.74	20.23	10.26	5.4	2.42	0.8468	0.8331	0.8215	0.8075	0.7825
65	-	92.1	34.21	13.59	4.57	0.8757	0.8634	0.8520	0.8380	0.8137
67	48.44	18.12	9.243	4.86	2.17	0.8162	0.8031	0.7917	0.7777	0.7528
68	-	-	18.22	8.882	3.6	0.8670	0.8541	0.8426	0.8280	0.8032
69	131.3	43.21	19.85	9.421	3.76	0.8518	0.8392	0.8280	0.8139	0.7893
74	59.11	22.04	11.26	5.86	2.59	0.8448	0.8313	0.8196	0.8054	0.7804
75	-	26.22	13.16	6.83	3.03	0.8512	0.8385	0.8270	0.8132	0.7884
76	89.99	30.35	14.38	7.05	2.95	0.8515	0.8373	0.8268	0.8118	0.7872
77	86.21	29.4	14	6.88	2.86	0.8502	0.8369	0.8256	0.8115	0.7867
78	80.54	27.83	13.33	6.58	2.76	0.8501	0.8370	0.8256	0.8113	0.7865
79	-	-	11.38	6.04	2.71	0.8680	0.8549	0.8434	0.8284	0.8034
80	-	-	12.52	6.31	2.7	0.8682	0.8548	0.8432	0.8277	0.8017
81	-	24.62	12.04	6.09	2.62	0.8667	0.8537	0.8419	0.8274	0.8016
82	-	23.7	11.68	5.94	2.55	0.8665	0.8532	0.8424	0.8270	0.8013
87	37.7	14.84	7.9	4.34	2.02	0.8699	0.8560	0.8441	0.8296	0.8040
88	70.05	24.78	12.05	6.02	2.56	0.8461	0.8327	0.8216	0.8068	0.7820
89	462.2	74.24	25.5	10.12	3.63	1.0217	1.0076	0.9952	0.9801	0.9537
90	-	-	-	39.4	8.585	0.9163	0.9033	0.8920	0.8778	0.8526
91	95.47	32.17	15.17	7.37	3.02	0.8474	0.8347	0.8236	0.8098	0.7852
99	-	-	-	5.5	2.67	0.8669	0.8544	0.8430	0.8287	0.8041
100	-	-	-	7.16	3.22	0.8447	0.8322	0.8208	0.8067	0.7819
101	-	-	10.83	5.86	2.72	0.8682	0.8550	0.8432	0.8285	0.8029
102	-	27.87	14.23	7.4	3.26	0.8501	0.8371	0.8256	0.8118	0.7868
103	-	-	12.03	6.15	2.7	0.8678	0.8548	0.8432	0.8283	0.8030
104	-	29.01	14.01	7	2.98	0.8510	0.8382	0.8268	0.8124	0.7874
106	-	-	-	5.11	2.4800	0.8172	0.8042	0.7927	0.7783	0.7530

107	-	31.18	15.62	7.933	3.4100	0.8278	0.8153	0.8040	0.7902	0.7658
108	-	-	14.92	7.67	3.32	0.8735	0.8610	0.8494	0.8354	0.8102
109	-	-	9.176	5.17	2.4800	0.8204	0.8076	0.7962	0.7823	0.7564
110	45.85	17.95	9.401	4.98	2.25	0.8422	0.8292	0.8178	0.8029	0.7775
111	79.08	27.84	13.55	6.77	2.87	0.8710	0.8579	0.8469	0.8324	0.8065
112	115.6	38.59	18.04	8.728	3.58	0.9133	0.9000	0.8873	0.8728	0.8475
113	176.5	51.41	22.46	10.13	3.87	0.9020	0.8887	0.8774	0.8635	0.8379
115	2187	326.6	92.25	28.4	7.27	0.9249	0.9118	0.9001	0.8860	0.8608
116	215.6	55.82	22.23	9.526	3.525	0.9295	0.9163	0.9044	0.8897	0.8642
117	-	-	-	6.31	2.94	0.8422	0.8292	0.8178	0.8037	0.7784
118	-	-	15.74	8.114	3.484	-	-	0.8473	0.8333	0.8086
119	224.3	47.54	18.78	8.36	3.29	1.0304	1.0159	1.0029	0.9880	0.9611
120	-	-	2.168@239F	6.154	2.813	-	-	0.8203@239F	0.8568	0.8313
121	-	-	2318	275.2	27.68	1.0387	1.0257	1.0142	0.9998	0.9744
122	-	-	798.4	128.4	17.94	0.9484	0.9356	0.9242	0.9099	0.8858
124	-	-	-	24.03	6.573	0.9890	0.9762	0.9648	0.9505	0.9256
125	-	-	41.02	15.73	5.132	0.9137	0.9014	0.8902	0.8764	0.8522
126	176.8	44.72	18.57	8.438	3.309	0.9723	0.9589	0.9469	0.9322	0.9062
127	1249	213.4	65.73	22.19	6.194	0.9108	0.8976	0.8866	0.8722	0.8478
128	137.8	40.4	18.01	8.443	3.408	0.9336	0.9205	0.9089	0.8945	0.8693
129	405.6	94.58	35.7	14.34	4.778	0.9055	0.8932	0.8816	0.8677	0.8426
130	331.5	75.45	28.76	11.92	4.183	0.9360	0.9230	0.9116	0.8971	0.8721
131	-	-	229.9	56.89	12.03	1.0421	1.0296	1.0185	1.0046	0.9803
132	-	540.6	147.8	45.46	11.51	0.9427	0.9310	0.9202	0.9071	0.8839
133	-	36.24	17.91	8.879	3.7600	0.8292	0.8168	0.8056	0.7916	0.7675
134	-	42.42	20.68	10.11	4.2650	0.8313	0.8189	0.8077	0.7938	0.7696
135	-	9.130@158F	6.337@185F	12.15	4.686	0.8657	0.8532	0.8420	0.8281	0.8036
136	-	63.22	28.16	12.89	4.886	0.8529	0.8406	0.8296	0.8158	0.7918
137	-	11.78@157F	8.054@185F	15.85	5.891	0.8659	0.8534	0.8183@185F	0.8283	0.8039
138	-	-	37.07	16.52	6.047	0.8542	0.8420	0.8312	0.8177	0.7940
139	668.6	125.5	41.19	14.76	4.494	0.8823	0.8698	0.8584	0.8449	0.8203
140	-	10.64@176F	8.00@194F	-	5.961	-	-	-	-	0.9292
141	-	108	39.14	15.45	5.119	0.9189	0.9066	0.8952	0.8812	0.8571
142	-	-	42.52	16.6	5.468	-	-	0.9474	0.9328	0.9077
143	-	77.75	32.9	14.4	5.233	-	0.9016	0.8905	0.8767	0.8526
144	-	-	4.527@239F	18.39	6.395	-	-	0.9149@239F	0.9509	0.9260
145	-	-	5.748@239F	2.971@275F	8.444	-	-	0.8614@239F	0.8486@275F	0.8718
146	-	3.609@239F	2.571@275F	13.93	4.907	-	0.9118@239F	0.8988@275F	0.9469	0.9218

152	-	-	10.21	5.671	2.677	0.8659	0.8530	0.8414	0.8271	0.8020
153	-	-	13.6	7.19	3.191	0.8458	0.8331	0.8217	0.8076	0.7827
155	-	-	43.83@176F	27.48@194F	19.09	0.9598	0.9482	0.9133@176F	0.9076@194F	0.9024
156	-	-	2.668@239F	8.924	3.58	-	-	0.8023@239F	0.8372	0.8125
157	-	29.28	14.49	7.435	3.268	-	0.8344	0.8232	0.8094	0.7856
158	-	-	2.183@239F	6.549	2.887	-	-	0.7966@239F	0.8328	0.8079
159	-	24.66	12.64	6.63	2.963	-	0.8325	0.8212	0.8074	0.7829
161	86.55	29.07	13.7	6.656	2.717	0.8664	0.8530	0.8413	0.8268	0.8009
162	95.74	31.15	14.44	6.918	2.809	0.8483	0.8355	0.8241	0.8099	0.7849
163	23.64	10.24	5.73	3.258	1.5840	0.8104	0.7976	0.7858	0.7712	0.7455
164	-	-	16.86	8.526	3.5950	0.8270	0.8146	0.8037	0.7897	0.7654
165	-	-	346.5	70.72	12.16	0.9926	0.9796	0.9682	0.9538	0.9287
166	-	-	290.6	60.38	10.93	0.9522	0.9395	0.9282	0.9140	0.8893
167	110.8	35.08	16.23	7.778	3.112	0.8745	0.8614	0.8495	0.8350	0.8093
168	-	49.87	21.48	9.555	3.555	0.8705	0.8574	0.8456	0.8312	0.8055
169	164.8	48.12	20.77	9.332	3.48	0.8518	0.8390	0.8279	0.8139	0.7892
170	303.3	63.76	24.96	10.68	3.95	1.0059	0.9922	0.9802	0.9654	0.9392
171	-	91.61	31.97	12.65	4.317	0.9960	0.9826	0.9707	0.9560	0.9302
172	-	-	110.9	33.75	8.371	0.9081	0.8957	0.8847	0.8712	0.8472
173	392.4	100.7	39.42	15.94	5.21	0.9008	0.8879	0.8765	0.8623	0.8373
174	-	-	17.56	8.384	2.882	0.9275	0.9145	0.9029	0.8885	0.8631
175	-	36.13	17.25	8.538	3.536	0.8806	0.8681	0.8568	0.8429	0.8184
176	-	-	4.908@158F	3.659@185F	2.8790	0.8196	0.8068	0.7748@158F	0.7651@185F	0.7565
177	592.3	132.2	47.79	18.16	5.583	0.9075	0.8946	0.8832	0.8692	0.8446
178	109.3	34.68	15.92	7.546	2.993	0.8705	0.8577	0.8461	0.8320	0.8069
179	-	7830	837.2	133.7	18.28	1.0372	1.0240	1.0122	0.9977	0.9722
180	67.99	24.34	11.95	6.13	2.633	0.8388	0.8260	0.8146	0.8007	0.7759
181	253.3	69.09	28.47	12.16	4.261	0.8928	0.8798	0.8684	0.8540	0.8290
182	2514	369.8	104.1	32.4	8.2630	0.8325	0.8204	0.8096	0.7961	0.7725
183	121.2	33.24	14.5	6.755	2.6920	0.8214	0.8088	0.7975	0.7835	0.7589
184	217.6	49.77	19.48	8.27	3.0140	0.8168	0.8043	0.7929	0.7789	0.7541
188	-	-	16.96	8.31	3.433	0.9019	0.8889	0.8773	0.8628	0.8376
189	-	-	10.36@158F	6.936@185F	5.073	0.9497	0.9369	0.9050@158F	0.8955@185F	0.8865
190	-	-	7.031@177.4F	5.689@196F	4.9040	0.8296	0.8171	0.7795@176F	0.7733@194F	0.7677
191	-	27.57	14.05	7.248	3.1240	0.8252	0.8125	0.8012	0.7873	0.7626
192	2931	406.7	112.1	34.13	8.429	0.9106	0.8982	0.8872	0.8734	0.8494
193	-	61.64	27.52	12.68	4.83	0.9076	0.8954	0.8842	0.8704	0.8463
196	2616	327	88.02	27.23	7.083	0.9554	0.9432	0.9323	0.9187	0.8951

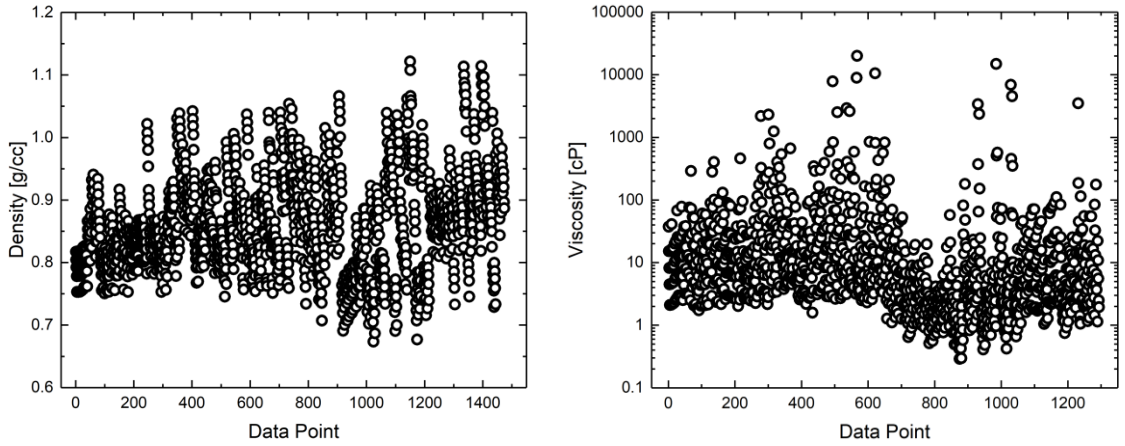
197	-	-	5.456@176F	4.493@194F	3.8240	0.8243	0.8119	0.7744@176F	0.7681@194F	0.7626
199	236.1	62.68	25.85	11.47	4.087	0.9025	0.8900	0.8789	0.8650	0.8408
200	-	-	586.9@176F	217.1@194F	106.1	1.0393	1.0277	0.9927@176F	0.9869@194F	0.9817
202	-	30.69	14.86	7.457	3.162	0.8774	0.8650	0.8538	0.8398	0.8155
203	-	-	-	20000000	8970	-	-	-	-	-
204	-	-	-	-	20000	-	-	-	-	-
205	-	-	5.433@239F	3.910@275F	7.4790	0.8377	0.8249	0.7642@239F	0.7514@275F	0.7745
206	-	-	3.296@239F	11.83	4.479	-	-	0.9285@239F	0.9658	0.9395
207	-	-	5.278@239F	3.595@275F	7.733	-	-	0.8390@239F	0.8265@275F	0.8494
208	42.83	17.75	9.631	5.317	2.496	0.8657	0.8526	0.8411	0.8267	0.8018
209	-	-	12.26	6.478	2.863	-	0.8317	0.8202	0.8061	0.7816
210	76.13	26.68	15.75	6.347	2.6290	0.8261	0.8133	0.8018	0.7878	0.7631
211	165.1	53.75	24.5	11.35	4.2210	0.8327	0.8205	0.8095	0.7957	0.7721
215	-	-	5.026	3.394	7.313	1.0400	1.0271	1.0154	1.0011	0.9759
216	-	3.886@239F	2.711@275F	16.56	5.512	0.9414	0.9294	0.9187	0.9050	0.8817
218	-	-	6.418@239F	4.114@275F	10.13	1.0080	0.9947	0.9827	0.9677	0.9418
219	851.8	157.1	52.7	18.89	5.864	0.9444	0.9320	0.9210	0.9070	0.8834
220	-	-	3.470@239F	2.569@275F	4.5940	0.8314	0.8183	0.8067	0.7922	0.7669
223	117.5	36.36	16.68	7.837	3.2040	0.8221	0.8092	0.7979	0.7837	0.7592
224	-	-	9.052@239F	5.406@275F	15.05	-	1.0542	0.9946@239F	0.9818@275F	1.0045
225	10500	815.9	170.7	41.98	9.121	0.9525	0.9398	0.9285	0.9150	0.8916
226	6.673@239F	4.253@275F	212.6	50.16	10.46	0.9958@239F	0.9833@275F	1.0458	1.0310	1.0059
228	5.272@239F	434.1	110.3	31.72	7.941	0.8810@239F	0.9387	0.9278	0.9140	0.8906
229	-	29.46	15.15	7.923	3.456	0.8722	0.8596	0.8482	0.8341	0.8096
230	4.992@239F	3.395@275F	74.56	25.25	7.256	0.9184@239F	0.9053@275F	0.9700	0.9547	0.9291
231	564.1	127.8	48.29	18.75	6.178	0.9291	0.9158	0.9046	0.8903	0.8662
232	-	-	831.8	140.4	19.9	0.9962	0.9837	0.9726	0.9587	0.9345
235	20.19	7.981	4.423	2.521	1.274	0.8731	0.8594	0.8473	0.8320	0.8055
236	3.105@239F	2.209@275F	33.25	12.79	4.316	0.884@239F	0.8754@275F	0.9393	0.9244	0.8989
237	209.9	55.38	22.99	9.921	3.732	0.9031	0.8905	0.8790	0.8650	0.8409
500	12.47	5.854	3.515	2.125	1.1080	0.8012	0.7878	0.7756	0.7606	0.7340
501	70.32	18.97	8.569	4.217	1.84	0.9774	0.9628	0.9497	0.9333	0.9046
502	22.63	9.155	4.966	2.771	1	0.8679	0.8541	0.8417	0.8262	0.7994
503	-	16.73	7.956	4.023	1.788	0.9647	0.9498	0.9375	0.9215	0.8939
504	30.41	11.6	6.058	3.274	1.523	0.8450	0.8317	0.8196	0.8048	0.7786
505	-	44.29	16.72	6.988	2.559	0.8962	0.8829	0.8714	0.8569	0.8315

506	17.16	7.428	4.225	2.486	1.266	0.8848	0.8705	0.8578	0.8420	0.8140
507	-	-	-	48.48	7.307	0.9615	0.9474	0.9350	0.9197	0.8927
509	52.8	19.09	9.534	4.884	2.142	0.8469	0.8337	0.8220	0.8076	0.7828
510	-	6.552	3.86	2.328	1.2360	0.8063	0.7932	0.7815	0.7665	0.7407
511	14.92	6.832	4.008	2.392	1.2430	0.8043	0.7909	0.7789	0.7640	0.7377
512	3.612	2.184	1.517	1.065	0.6449	0.7772	0.7634	0.7510	0.7352	0.7072
513	4.653	2.64	1.794	1.219	0.7328	0.8711	0.8571	0.8440	0.8277	0.7988
514	6.749	3.677	2.407	1.591	0.9086	0.8373	0.8235	0.8110	0.7956	0.7684
516	8.983	4.484	2.844	1.84	1.075	1.0152	0.9998	0.9860	0.9690	0.9387
517	15.98	6.87	3.928	2.358	1.24	0.9498	0.9356	0.9231	0.9071	0.8790
518	19.33	8.494	4.896	2.896	1.476	0.9069	0.8932	0.8808	0.8658	0.8389
519	-	-	-	1.808	1.06	1.0042	0.9890	0.9752	0.9583	0.9289
520	-	7.951	4.62	2.725	1.422	0.8880	0.8740	0.8616	0.8466	0.8200
521	9.125	4.709	2.935	1.89	1.104	0.9409	0.9262	0.9133	0.8974	0.8693
522	6.24	3.606	2.388	1.671	0.9911	0.9321	0.9175	0.9043	0.8889	0.8607
523	9.609	4.977	3.222	2.109	1.193	0.8810	0.8675	0.8556	0.8407	0.8136
524	-	-	2.201	1.51	0.9151	-	-	0.9929	0.9756	0.9453
525	11.71	5.685	3.515	2.187	1.81	0.8907	0.8768	0.8644	0.8488	0.8216
526	-	-	-	19.72	4.616	1.0662	1.0530	1.0412	1.0266	1.0007
527	-	-	-	10.22	3.301	0.9508	0.9383	0.9268	0.9124	0.8877
528	2.265	1.492	1.102	0.8026	0.5156	0.7636	0.7487	0.7360	0.7194	0.6907
529	2.948	1.864	1.345	0.9594	0.6010	0.7704	0.7563	0.7436	0.7279	0.7004
530	2.551	1.663	1.218	0.8776	0.5597	0.7798	0.7656	0.7528	0.7366	0.7081
531	-	2.323	1.627	1.137	0.6940	0.7771	0.7629	0.7505	0.7347	0.7071
532	-	2.842	1.953	1.335	0.7960	0.7826	0.7685	0.7562	0.7408	0.7135
533	4.16	2.519	1.762	1.224	0.7454	0.7898	0.7757	0.7633	0.7478	0.7209
534	-	3.453	2.331	1.559	0.9054	0.7874	0.7737	0.7612	0.7460	0.7190
535	-	-	2.742	1.795	1.0240	0.7916	0.7780	0.7658	0.7506	0.7240
536	-	3.729	2.496	1.67	0.9647	0.7999	0.7861	0.7737	0.7585	0.7314
537	-	-	3.209	2.062	1.1500	0.7964	0.7828	0.7708	0.7558	0.7292
538	4.288	2.575	1.802	1.254	0.7648	0.8712	0.8565	0.8442	0.8275	0.7988
539	6.311	3.515	2.351	1.576	0.9119	0.8282	0.8139	0.8017	0.7864	0.7590
540	-	-	4.29	2.664	1.4240	0.8021	0.7888	0.7729	0.7621	0.7361
541	-	-	-	4.191	2.0830	0.8118	0.7988	0.7873	0.7728	0.7475
542	17.62	7.752	4.432	2.589	1.302	0.8361	0.8228	0.8107	0.7958	0.7693
543	2.83	1.853	1.346	0.9569	-	0.8863	0.8695	0.8543	0.8353	-
544	57.78	21	10.66	5.589	2.49	0.8853	0.8725	0.8611	0.8469	0.8220
545	6.029	3.178	2.052	1.335	0.7598	0.7898	0.7761	0.7635	0.7482	0.7210
546	2.768	1.64	1.146	0.801	0.4940	0.7751	0.7606	0.7478	0.7315	0.7020
547	2.302	1.456	1.055	0.762	0.4820	0.7865	0.7718	0.7588	0.7424	0.7128

548	7.054	4.273	2.976	2.05	1.223	0.9229	0.9097	0.8979	0.8833	0.8576
549	0.9413	0.6822	0.5412	0.4184	0.2900	0.7507	0.7354	0.7219	0.7048	0.6734
550	0.9274	0.6842	0.5455	0.4268	0.2982	0.7662	0.7505	0.7365	0.7189	0.6866
551	1.985	1.434	1.12	0.8566	0.5775	0.8793	0.8646	0.8514	0.8349	0.8052
552	81.99	28.25	13.91	6.976	2.936	0.9185	0.9053	0.8935	0.8789	0.8532
553	180.3	48.27	20.19	8.935	3.345	0.9090	0.8950	0.8841	0.8696	0.8440
554	17.52	7.846	4.516	2.642	1.3220	0.8059	0.7926	0.7808	0.7662	0.7401
555	5.259	3.045	2.081	1.413	0.8280	0.8152	0.8013	0.7886	0.7727	0.7443
556	8.434	4.584	2.985	1.938	1.0750	0.8136	0.8004	0.7886	0.7738	0.7470
557	9.689	4.524	2.742	1.697	0.9176	0.8050	0.7879	0.7758	0.7608	0.7338
559	-	23.18	11.11	5.595	2.423	0.9412	0.9279	0.9159	0.9010	0.8750
560	20.22	7.36	4.012	2.325	1.1920	1.0394	1.0235	1.0093	0.9916	0.9607
561	12.3	6.559	4.218	2.706	1.482	0.9624	0.9489	0.9367	0.9216	0.8949
562	-	3387	374.8	65.38	11.21	0.9785	0.9661	0.9550	0.9412	0.9171
563	-	-	2366	152.4	15.32	0.9891	0.9765	0.9654	0.9513	0.9269
564	24.53	10.11	5.628	3.25	1.635	0.9500	0.9366	0.9247	0.9099	0.8838
565	1.507	1.054	0.8085	0.6118	0.4080	0.7653	0.7505	0.7372	0.7205	0.6903
566	1.976	1.329	0.9989	0.7389	0.4810	0.7730	0.7586	0.7455	0.7292	0.7000
567	6.132	3.351	2.226	1.494	0.8817	1.0357	1.0211	1.0079	0.9914	0.9624
568	-	-	1.72	1.197	0.738	1.0215	1.0063	0.9927	0.9758	0.9461
569	5.558	3.349	2.31	1.569	0.9162	0.9119	0.8968	0.8834	0.8665	0.8369
570	3.214	2.109	1.546	1.114	0.696	0.8845	0.8698	0.8565	0.8399	0.8106
571	6.847	3.81	2.535	1.694	0.987	0.8693	0.8554	0.8430	0.8275	0.8001
572	10.25	5.331	3.386	2.16	1.188	0.8319	0.8186	0.8067	0.7918	0.7655
573	6.203	3.59	2.434	1.641	0.9606	0.8246	0.8109	0.7986	0.7834	0.7563
574	-	14900	512.7	54.28	8.15	1.0620	1.0492	1.0378	1.0236	0.9987
575	571.9	75.06	25.51	10.38	3.792	0.9927	0.9803	0.9692	0.9557	0.9320
576	-	-	6.162@176F	4.684@194F	3.776	1.1217	1.1078	1.0662@176F	1.0593@194F	1.0531
577	49.76	17.81	9.31	5.086	2.379	0.9950	0.9822	0.9706	0.9563	0.9313
578	72.45	24.4	12.11	6.278	2.772	1.0019	0.9890	0.9775	0.9634	0.9386
580	3.558	2.334	1.706	1.231	0.7823	0.8924	0.8782	0.8657	0.8497	0.8218
581	1.638	1.124	0.8513	0.6377	0.4210	0.7513	0.7367	0.7235	0.7070	0.6769
582	6.109	3.437	2.308	1.541	0.8949	0.7845	0.7707	0.7586	0.7434	0.7162
583	-	5.061	3.235	2.075	1.1480	-	0.7795	0.7675	0.7527	0.7262
584	-	-	7.454	4.308	2.1260	-	-	0.7858	0.7713	0.7460
586	-	6918	454.8	62.54	10.08	0.9833	0.9709	0.9596	0.9455	0.9213
587	-	4512	351	56.13	10.13	1.0200	0.9981	0.9873	0.9738	0.9503
588	-	8.302	5.023	3.051	1.5890	-	0.7934	0.7815	0.7668	0.7412
589	3.348	2.097	1.502	1.064	0.6590	0.7861	0.7720	0.7594	0.7435	0.7155
590	-	3.082	2.109	1.436	0.8515	0.7953	0.7813	0.7691	0.7537	0.7265

591	16.49	7.89	4.788	2.836	1.4800	0.8053	0.7921	0.7801	0.7654	0.7394
592	9.55	4.793	3.042	1.937	1.084	0.9436	0.9294	0.9168	0.9011	0.8730
593	40.78	15.76	8.394	4.598	2.163	0.9158	0.9025	0.8908	0.8761	0.8502
594	46.14	18.04	9.57	5.147	2.359	0.8811	0.8679	0.8563	0.8420	0.8167
595	20.896	9.192	5.34	3.13	1.609	0.9077	0.8941	0.8820	0.8667	0.8402
596	32.166	13.572	7.486	4.243	2.038	0.8795	0.8650	0.8540	0.8393	0.8136
597	24.186	10.599	6.044	3.529	1.779	0.9104	0.8967	0.8846	0.8695	0.8426
598	29.403	12.746	7.113	4.052	2.055	0.8676	0.8631	0.8510	0.8361	0.8100
599	56.83	18.48	8.71	4.34	1.86	0.9261	0.9103	0.8997	0.8844	0.8574
600	71.383	20.843	9.276	4.396	1.81	0.9108	0.8967	0.8843	0.8690	0.8418
601	70.973	21.151	9.459	4.551	1.928	0.8817	0.8683	0.8566	0.8418	0.8157
602	25.825	10.81	6.033	3.379	1.697	0.9057	0.8918	0.8794	0.8642	0.8373
603	33.5	13.32	7.164	3.93	1.841	0.8723	0.8588	0.8470	0.8323	0.8063
604	41.07	15.365	7.917	4.186	1.884	0.9150	0.9011	0.8888	0.8736	0.8466
605	45.7	15.77	7.784	4.014	1.776	0.8829	0.8694	0.8574	0.8425	0.8165
606	9.459	4.66	2.932	1.861	1.045	0.9794	0.9651	0.9521	0.9363	0.9083
607	11.01	5.444	3.4	2.122	1.15	0.8899	0.8760	0.8639	0.8487	0.8219
608	-	4.073	2.678	1.763	-	-	0.8865	0.8739	0.8581	0.8421
609	-	-	3.708	2.36	1.2810	0.7989	0.7857	0.7737	0.7590	0.7328
610	65.32	22.03	10.54	5.257	2.227	0.9424	0.9284	0.9164	0.9014	0.8748
611	61.43	20.81	10.08	5.017	2.158	0.9173	0.9039	0.8916	0.8766	0.8500
612	74.55	24.03	11.24	5.491	2.283	0.8843	0.8713	0.8595	0.8451	0.8196
613	81.31	24.89	11.49	5.506	2.271	0.9470	0.9330	0.9209	0.9058	0.8796
614	92.05	29.21	13.3	6.287	2.5	0.9222	0.9086	0.8967	0.8818	0.8555
615	111.17	30.76	13.17	5.977	2.346	0.8888	0.8755	0.8639	0.8494	0.8239
616	-	-	3.542@239F	2.431@275F	5.154	1.1134	1.0998	1.0877	1.0727	1.0461
617	-	3.119@239F	2.151@275F	17.24	4.473	1.0808	1.0684	1.0550	1.0384	1.0094
618	84.82	25.56	12.1	6.123	2.659	0.9684	0.9553	0.9438	0.9296	0.9046
619	-	1.501@239F	1.165@275F	3.759	1.9260	0.8098	0.7968	0.7852	0.7707	0.7454
620	3.322	2.208	1.619	1.165	0.749	0.8934	0.8782	0.8646	0.8475	0.8177
622	5.079	3.144	2.212	1.534	0.9186	0.9031	0.8889	0.8763	0.8606	0.8329
623	8.001	4.507	3.009	1.992	1.138	0.9170	0.9031	0.8909	0.8757	0.8489
625	-	17.34	8.227	4.26	1.945	1.0373	1.0241	1.0123	0.9973	0.9709
626	15.51	7.591	4.729	2.872	1.582	0.9582	0.9448	0.9336	0.9187	0.8931
627	-	-	13.46	6.189	2.488	0.9488	0.9352	0.9231	0.9079	0.8815
628	-	25.42	11.64	5.593	2.333	0.9169	0.9035	0.8916	0.8770	0.8514
629	-	20.78	10.07	5.08	2.213	0.8825	0.8697	0.8582	0.8439	0.8186
631	-	-	2.741@239F	1.903@275F	4.004	1.1141	1.0984	1.0845	1.0668	1.0359
632	-	-	2.859@239F	14.22	4.128	-	-	0.8742@239F	0.9103	0.8853

633	-	-	2.932@239 F	24.19	4.612	1.0972	1.0814	1.0672	1.0496	1.0185
634	3503.2	186.8	45.1	14.17	4.323	0.9533	0.9404	0.9290	0.9146	0.8894
637	-	-	1.140@239 F	1.059@27 5F	1.419	0.9450	0.9308	0.9182	0.9024	0.8748
638	114.8	31.42	13.71	6.303	2.58	0.9639	0.9507	0.9389	0.9243	0.8991
639	2.022@23 9F	34.36	14.68	6.701	2.688	0.8593@23 9F	0.9217	0.9098	0.8951	0.8695
640	55.23	19.03	9.496	4.881	2.184	0.8884	0.8755	0.8641	0.8499	0.8250
641	-	-	-	3.299@23 9F	2.296@27 5F	-	-	-	1.0395@23 9F	1.0261@27 5F
642	13.59	6.207	3.686	2.218	1.1630	0.7956	0.7821	0.7703	0.7551	0.7289
643	17.1	7.525	4.369	2.567	1.3180	0.7996	0.7861	0.7742	0.7591	0.7333
645	54.77	18.37	9.324	4.934	2.277	0.9729	0.9594	0.9474	0.9325	0.9070
646	54.88	14.76	6.913	3.533	1.632	1.0065	0.9923	0.9798	0.9639	0.9363
647	16.6	7.634	4.57	2.752	1.446	0.9108	0.8973	0.8854	0.8701	0.8439
648	-	84.24	32.91	13.21	4.517	0.9447	0.9316	0.9200	0.9053	0.8800
652	176.2	6.57	3.688	2.16	1.143	0.9827	0.9683	0.9554	0.9392	0.9114
654	1.937@23 9F	22.26	10.86	5.554	2.5	0.8869@23 9F	0.9455	0.9346	0.9206	0.8967



(a) (b)

Figure C.1 Experimental (a) density and (b) dynamic viscosity of all 305 hydrocarbon molecules from API Project 42 [263].

C.3. Models with Combined Static and Dynamic Descriptors

The best 2 and 3 static and dynamic descriptors are Equations C.1 and C.2. Model I is a function of the simulation-calculated density (*rho*) and Ghose-Crippen LogKow (*ALogP*) [259]. Model III is a function of *rho*, *ALogP*, and

pairwise interaction energy ($epair$). $epair$ quantify strength of inter and intra-molecular interactions.

Model I: $$\rho(T) = f(ALogP, rho) \quad \text{(C.1)}$$

Model II: $$\rho(T) = f(ALogP, rho, epair) \quad \text{(C.2)}$$

The best 2, 3, and 4 static and dynamic descriptors are Equations C.3, C.4, and C.5. Model I is a function of complexity of a system ($fragC$), bonding energy ($ebond$) of the molecule. Model III is a function of kinetic energy, number of rotatable bonds ($nRotB$) in the molecule, and lipophilicity index ($XLogP$). Model III is a function of temperature (T), fraction of rotatable bond ($RotBFrac$), excluding terminal bonds in a molecule, and Crippen's molar refractivity ($CrippenMR$), and sphericity (b) of molecules. b quantify how spherical a molecule is. If b is 0 then a molecule is perfectly spherical in shape. A higher b mean a molecule is further away from the spherical shape or distribution of its atoms. $ALogP$ and $XLogP$ are the lipophilicity indexes which were calculated using different approaches. Lipophilicity indexes quantify the interaction between molecules of a system.

Model I: $$\log\eta(T) = f(fragC, ebond) \quad \text{(C.3)}$$

Model II: $$\log\eta(T) = f(ke, nRotB, XLogP) \quad \text{(C.4)}$$

Model III: $$\log\eta(T) = f(T, RotBFrac, CrippenMR, b) \quad \text{(C.5)}$$

Table C.3 lists the R^2 and RMSE values for the density and viscosity models with combined static and dynamic descriptors for the training, validation, and test data sets. The average R^2 value of the training, validation, and test data sets is also listed for the comparison between the models. The maximum value of VIF out of VIFs for each predictor is also listed to indicate the degree of multicollinearity.

Table C.3 Model assessment parameters for the density and viscosity models with combined static and dynamic descriptors.

Parameter		Density Models		Viscosity Models		
		Model I	Model II	Model I	Model II	Model III
Training	R^2	0.999	1.000	0.998	0.992	1.000
	RMSE	0.003	0.001	0.028	0.066	0.009
Validation	R^2	0.999	1.000	1.000	0.992	1.000
	RMSE	0.003	0.000	0.012	0.065	0.000

Test	R^2	0.982	0.990	0.818	0.947	0.947
	RMSE	0.011	0.008	0.288	0.155	0.156
Average R^2		0.993	0.997	0.939	0.977	0.982
Maximum VIF		0.999	1.000	1.109	4.094	2.407

Figures C.2(a) and C.2(b) show density and viscosity predicted using their best models of combined static and dynamic descriptors for all data sets over a wide temperature range. The blue dashed line represents an ideal prediction. The model-predicted density for the training, validation, and test data sets is shown as black circles, red squares, and green triangle, respectively.

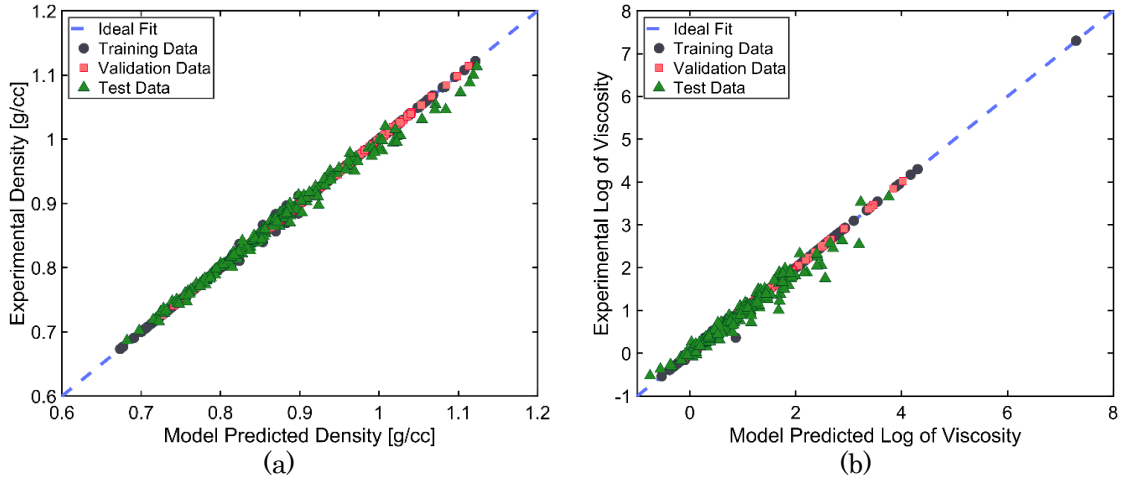


Figure C.2 Model predicted (a) density (Equation C.2) and (b) viscosity (Equation C.5) by the models with combined static and dynamic descriptors for the training (black circle), validation (red square), and test data (green triangle) sets. The blue dash lines represent the ideal prediction.

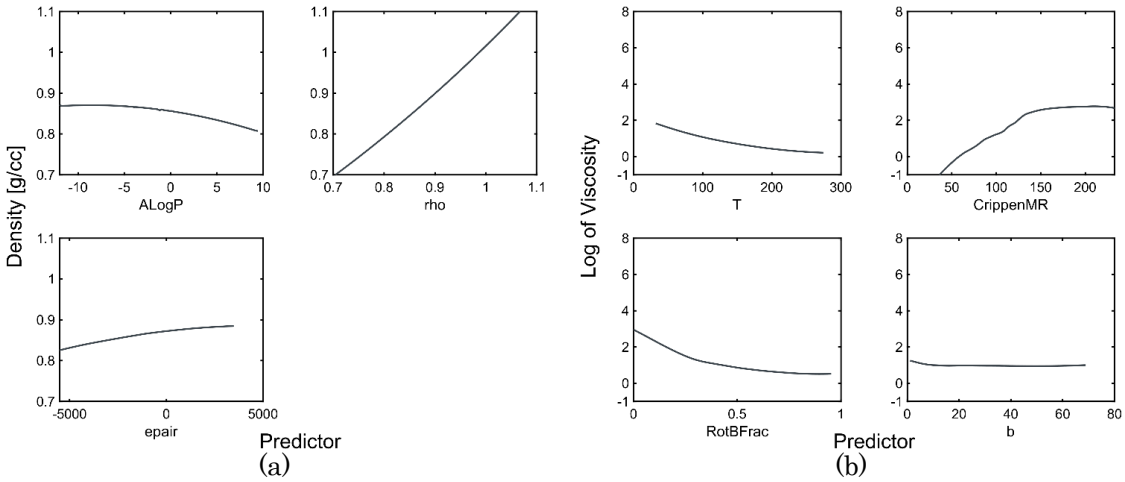


Figure C.3 Partial dependency plots of predictors in the best (a) density model (Equation C.2) and (b) viscosity model (Equation C.5) with combined static and dynamic descriptors.

C.4. Density Models with Dynamic Descriptors excluding Simulation-Calculated Density

The best three models with 1, 2, and 3 dynamic descriptors excluding simulation-calculated density (rho) for predicting the temperature-dependent density of the hydrocarbons are Equations C.6, C.7, and C.8. Model I is only a function of van der Waals energy long-range tail correction ($etail$). Models II is a function of $etail$ and energy due to improper interaction ($eimp$, higher for more rigidity molecule). Model III has an additional term which is the length of the molecule (L_y). It was noted that length of the molecules in any direction (L_x, L_y, L_z) resulted in similar predicting performance when used in Equation C.8, likely because the dynamic descriptors were calculated from equilibration MD simulations with a cubic simulation box.

Model I $$\rho(T) = f(etail) \tag{C.6}$$

Model II $$\rho(T) = f(etail, eimp) \tag{C.7}$$

Model III $$\rho(T) = f\left(etail, eimp, L_y\right) \tag{C.8}$$

Table C.4 lists the R^2 and RMSE values for the density models with dynamic descriptors excluding simulation-calculated density for the training, validation, and test data sets. The average R^2 value of the training, validation, and test data sets is also listed for the comparison between the models. The maximum value of VIF out of VIFs for each predictor is also listed to indicate the degree of multicollinearity.

Table C.4 Model assessment parameters for the density models with dynamic descriptors excluding simulation-calculated density.

Parameter		Model I	Model II	Model III
Training	R^2	0.928	0.979	1.000
	RMSE	0.020	0.011	0.001
Validation	R^2	0.913	0.989	1.000
	RMSE	0.024	0.008	0.000
Test	R^2	0.929	0.970	0.973
	RMSE	0.022	0.014	0.013
Average R^2		0.923	0.980	0.991
Maximum VIF		1.000	3.246	3.336

Figures C.4(a) shows density predicted using the best models of dynamic descriptors excluding simulation-calculated density for all data sets over a wide temperature range. The blue dashed line represents an ideal prediction. The model-predicted density for the training, validation, and test data sets is shown as black circles, red squares, and green triangle, respectively. Figures C.4(b) shows partial dependency plots of predictors in the best density model predicted of dynamic descriptors excluding simulation-calculated density

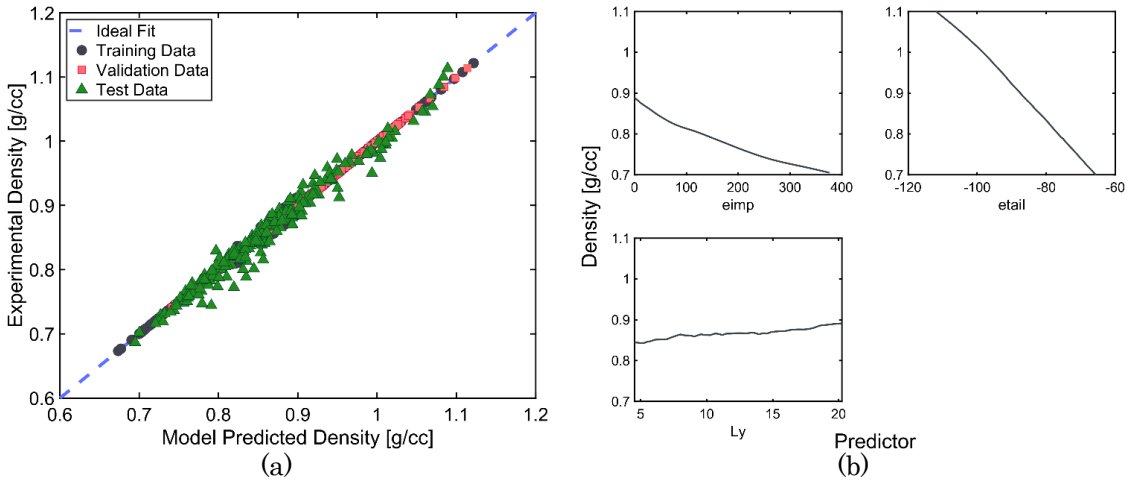


Figure C.4 (a) Model predicted density (Equation C.8) by the models with dynamic descriptors for the training (black circle), validation (red square), and test data (green triangle) sets. The blue dash lines represent the ideal prediction. (b) Partial dependency plots of predictors in the best density model (Equation C.8) with dynamic descriptors.

C.5. Additional Supplementary Materials

The following supplementary materials are provided on GitHub at https://github.com/panwarp/Supplementary-Materials which include:

- Schematic of all the molecules
- Definition of important molecular descriptors.
- All experimental data with the important static and dynamic descriptors of all the molecules
- LAMMPS data files of all the molecules and input files to run the MD simulations
- MATLAB files of the best models to predict temperature dependent density and viscosity.