Journal of Molecular Liquids 190 (2014) 223–230

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

A group contribution method for determination of thermal conductivity of liquid chemicals at atmospheric pressure

Farhad Gharagheizi [a,b], Poorandokht Ilani-Kashkouli [a,b], Mehdi Sattari [a], Amir H. Mohammadi [a,c,**], Deresh Ramjugernath [a,*]

[a] Thermodynamics Research Unit, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa
[b] Department of Chemical Engineering, Buinzahra Branch, Islamic Azad University, Buinzahra, Iran
[c] Institut de Recherche en Génie Chimique et Pétrolier (IRGCP), Paris Cedex, France

ARTICLE INFO

Article history:
Received 27 May 2013
Received in revised form 17 September 2013
Accepted 22 October 2013
Available online 7 November 2013

Keywords:
Liquid thermal conductivity
Group contribution
Chemical structure
Database
Reliable model
Prediction

ABSTRACT

In this communication, a group contribution method (GC) for the representation/prediction of liquid thermal conductivity of pure chemical compounds, most of which are organic in nature, is presented. Nearly 19,000 liquid thermal conductivity data at different temperatures compiled for 1635 chemical compounds were extracted from the DIPPR 801 database and used to develop the proposed model, as well as to validate and optimize its parameters and evaluate its predictive capability. The parameters of the model comprise the occurrences/existence of 49 chemical substructures plus temperature. Nearly 80% of the data set (15,450 data points) is used to develop the model parameters, 10% of the data set (1931 data points) was employed to validate and optimize the model parameters, and the remaining data (1931 data points) were implemented to assess its predictive capability. The average absolute relative deviation of the model results with respect to the DIPPR 801 data is less than 7.1%. In terms of its simplicity and wide range of applicability, the model shows reasonable accuracy.

1. Introduction

In most chemical processes, there is a need to perform an energy balance on units, in order to determine how much heat is either given off or absorbed in various equipment such as furnaces, evaporators, distillation units, dryers, reaction vessels, etc. [1]. Heat transfer can occur by three mechanisms, one of which is thermal conduction. In thermal conduction, heat transfer occurs due to a temperature gradient [1]. It is attributed to vibrational movements of molecules and expressed as Fourier's Law [2].

Liquid thermal conductivity is an important transport property which is needed for calculation of the thermal conduction component of heat transfer. Accurate measurements of thermal conductivity are not straightforward and special care is needed in experimental measurements because of the possible presence of convective currents and heat losses. As a result, the uncertainty in the reported measurements of existing experimental data is relatively large compared to other thermophysical properties reported [3–7].

The aspects of kinetic theories of thermal conductivity for monoatomic liquids were initially presented in 1950 [8]. Since then, many researchers have attempted to develop models for the estimation of liquid thermal conductivities of pure compounds. Most of them are empirical in nature and have limited applicability. According to Sastri and Rao [9,10], in most of the existing methods, thermal conductivity is correlated at a reference temperature such as the normal boiling point (for instance in the models proposed by Sato-Riedel [7,11], and Sastri and Rao [9,10]) or at 293.15 K (for instance in the models proposed by Missenard [12] and API Technical Data Book [13]). Thus, the temperature dependency of the model is investigated.

A detailed review of the existing models for the estimation of liquid thermal conductivity of pure compounds has been presented by Poling et al. [14]. In this review [14], it has been mentioned that although the methods presented by Latini et al. and Baroncini et al. [15–21] and Sastri et al. [22] are generally better than the others below the normal boiling temperature, the deviation of the models vary widely, typically less than 15%. Additionally, although the Latini et al. and Baroncini et al. [15–21] methods have been successfully applied for refrigerants up to reduced temperatures equal 0.9, it has been mentioned that there are few experimental liquid thermal conductivity data for reduced temperatures greater than 0.65. Therefore, it can be concluded that the models currently available may not accurately predict the liquid thermal conductivity for reduced temperatures greater than 0.65. Moreover, the model proposed by Latini et al. and Baroncini et al. [15–21] is presented for several particular chemical families of compounds, viz. saturated hydrocarbons, olefins, cycloparafins, aromatics, alcohols, organic acids, ketones, esters, and refrigerants. These chemical families do not cover a significant

* Correspondence author at: Thermodynamics Research Unit, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa.
** Corresponding author.
E-mail addresses: a.h.m@irgcp.fr (A.H. Mohammadi), ramjuger@ukzn.ac.za (D. Ramjugernath).

0167-7322/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.molliq.2013.10.019

number of widely used chemical families such as amines, silanes/siloxanes, inorganic compounds, sulfides/thiophenes, mercaptanes, epoxides, peroxides, nitriles, elements, and aldehydes. This issue significantly affects its applicability domain when using the model. As a result, the model is not a general model.

Furthermore, the model proposed by Sastri et al. [22] is based on 748 data points for 208 compounds below the normal boiling point, and 186 data points for just 23 pure compounds above the normal boiling point. Therefore, it affirms that the model cannot be considered as a general model. Moreover, Poling et al. [14] mentioned that none of the existing methods can predict the large changes of thermal conductivity near the critical point.

As a result, there is a requirement to utilize the extensive databases of experimental liquid thermal conductivity data such as that available in DIPPR 801 [23] to develop more accurate and more comprehensive models. In this study, a group contribution method (GC) is presented for the representation/prediction of the liquid thermal conductivity of more than 1600 pure compounds (mostly non-electrolyte organic molecules) at different temperatures and atmospheric pressure for temperatures below the normal boiling point and at saturation pressure for temperatures above the normal boiling point. It should be noted that the normal boiling point is the boiling point at which the vapor pressure equals 1 atm.

2. The database

The DIPPR 801 database [23] was used to provide the liquid thermal conductivity data for this study. Nearly 19,000 liquid thermal conductivity data for 1635 pure compounds which are mostly organic molecules were extracted from the DIPPR 801 database [23] and were used to develop and then to validate the model. It is worth to mention that DIPPR 801 data which have been evaluated and labeled as "Accepted" data in the database were used in this study to develop and validate the model. It should be stated that the data is related to liquid thermal conductivity of pure chemical compounds at a pressure of 1 atm for temperatures below normal boiling point (normal boiling point is the boiling point at which the vapor pressure equals 1 atm) and at saturation pressure for temperatures above the normal boiling point.

In order to obtain a predictive model, it is required to split the data set into three sub-data sets; first one for developing the model (called the "training set"), the second one for assessing the internal validity of the model (called the "validation set"), and the last one for evaluating the predictive capability of the obtained model (called the "test set"). This can be done randomly; however, this may bring about some inappropriate allocation of data to each sub-data set. In order to avoid this issue, the K-means clustering technique can be used [24; 25]. This method partitions a data set into n sub-data sets in which each data point belongs to the subset with the closest mean. This procedure resolves the issue of inappropriate allocation of data sets. Another important point is the quota of each sub-data set from the main data set. It has been shown that if the training set is too small, the model produced does not have adequate predictive power. Moreover, if the data set is too large, the model may produce significantly better results for training set rather than for the validation and test sets [26]. In order to prevent these issues, nearly 80% of the data set was allocated to the training set (15,450 data points) and the remaining data was divided into two equal sub-data sets and were allocated to the validation and test sets, respectively (1931 data points each).

3. Development of a new group contribution model

In order to develop a group contribution model, the chemical structures of all 1635 pure compounds were analyzed with great attention paid to comparing the chemical substructures to determine the most efficient contributions to the liquid thermal conductivity. Initially, nearly 650 chemical substructures that have already been used by the authors in their previous studies plus temperature were considered.

Consequently, the occurrences of these 650 chemical substructures in each of the 1635 compounds were counted. The results were imported into a table. Then the pair correlation between each pair of these 650 chemical substructures was calculated to determine those chemical substructures that were linearly related to each other. In the next step, one of each pair of the chemical substructures that have a squared correlation coefficient higher than 0.95 was eliminated and the other one kept for the next step. The collection of chemical substructures diminished to 321 chemical substructures using the pair correlation technique.

The final step in developing the model is the selection of an optimal subset of chemical substructures which have the highest contribution for the liquid thermal conductivity of pure compounds. To do this a subset variable selection technique should be implemented to select the most statistically effective chemical substructures for the liquid thermal conductivity. This step should be done in order to avoid entering irrelevant chemical substructures that have no significant effect on liquid thermal conductivity. For this purpose, the sequential search algorithm in the MATLAB software platform was employed. The major target of a sequential search is to find an optimal subset of chemical substructures for a specified model size and develop a linear correlation using them. The basic idea of the method is to replace each chemical substructure one at a time with all the remaining ones and see whether an improved model is obtained. The major steps of the algorithm are as follows:

step 1 Introduce all the molecular descriptors
step 2 Consider all the one-variable linear correlations between the thermal conductivities of liquids and the molecular descriptors
step 3 The first optimal descriptor is the one which predicts the thermal conductivities of liquids with lowest possible AARD%

$$\text{AARD}\% = \frac{100}{N}\sum_{i}^{N}\frac{\left|k^{rep/pred} - k^{exp}\right|}{k^{exp}} \quad (1)$$

step 4 Define the number of molecular descriptor(s) in the first step $i = 1$
step 5 Consider all the linear correlations between the thermal conductivities of liquids and $i + 1$ molecular descriptors including the i optimal descriptor(s) selected in the previous step(s)

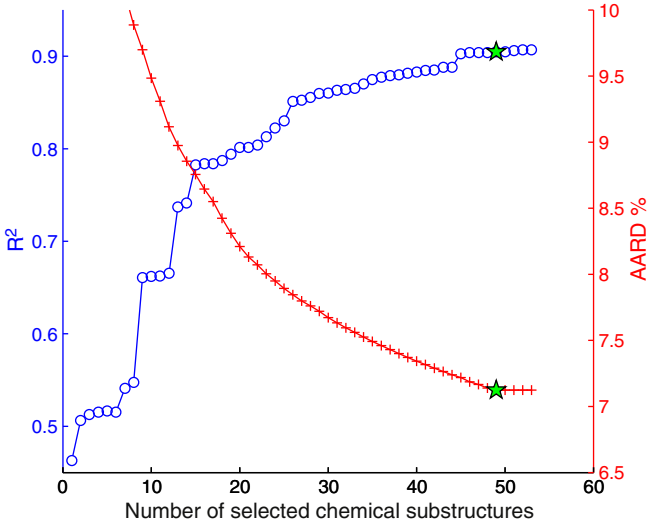

Fig. 1. The improvement of model in terms of R^2 and AARD% as function of number of chemical substructures selected by the program.