



A novel group contribution-based method for estimation of flash points of ester compounds



Dai Yimin*, Liu Hui, Li Xun, Zhu Zhiping, Zhang Yuefei*, Cao Zhong, Zhu Lixuan, Zhou Yue

Human Provincial Key Laboratory of Materials Protection for Electric Power and Transportation, School of Chemistry and Biological Engineering, Changsha University of Science and Technology, Changsha 410004, China

ARTICLE INFO

Article history:

Received 12 March 2014
Received in revised form 5 May 2014
Accepted 24 May 2014
Available online 2 June 2014

Keywords:

Flash point
Ester
Group contribution-based method
Multiple linear regression
Model validation
Quantitative structure–property relationship

ABSTRACT

Flammability in esters is one of the most important features for preparation, storage, safe processing, handling and shipping of a substance. In this study, a new accurate group contribution-based method was presented for estimation of the flash points of ester compounds. The predicted flash points for a dataset of 80 esters were in good agreement with the experimental values. The obtained results showed the squared correlation coefficient (R^2) of 0.9902, root mean square error (RMS) of 5.371K, and average absolute relative deviation (AARD) of 1.22% for the experimental values. To propose a predictive model, 30 ester compounds in testing set were investigated. The average percent error of 1.96% for the predicted flash point of the investigated compounds was found from the corresponding experimental values. The established model was validated and tested through the use of leave-one-out cross validation method, Y-randomization and applicability domain analysis. The results demonstrated the improved accuracy of the presented method with respect to previously proposed methods in open literatures. Therefore, the model can be well used to predict the flash points of a wide range of ester compounds, which is an accurate method for potential application.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

As the technology progresses, more and more flammable materials are operated in process industries. The flash point (FP) is one of the most significant properties of flammable liquids in industrial processes when evaluating process safety [1]. The flash point is defined as the lowest temperature at which it can vaporize to shape a flammable mixture in air. The combustible substance reacts with oxygen in the air in an exothermic oxidation reaction giving a momentary flash [2]. Obviously, the flash points of compounds are important in terms of both practical uses (i.e., combustion chemistry) and safety (i.e., handling, transporting and storing of the compounds in bulk quantities). Although the measured flash point values are available for many organic compounds by the open cup or closed-cup methods [3], experimental flash point values are scarce in literatures. And determining the flash point of a chemical by an experimental approach is very laborious and is not always feasible [4]. Hence it is necessary to develop theoretical prediction methods which are desirably convenient and reliable for predicting flash point.

Recently, a variety of quantitative structure–property relationship methods have been used to determine the flash points of different

compounds [5–9]. QSPR models utilizing theoretical molecular descriptors have the advantage of not needing data measurements for prediction of the property over several decades [10–15]. Stefanis et al. [10] developed a group contribution method that used first-order groups and second-order groups. The results had $R^2 = 0.967$, $S = 14.7K$, and a mean error of 3.27% for flash points including complex aromatic, multi-ring, and heterocyclic compounds. Katritzky et al. [11] published an update of their previous QSPR study of flash points using an extended dataset of 758 organic compounds by using geometrical, topological, quantum mechanical and electronic descriptors by linear and nonlinear methods. Keshavarz and Ghanbarzadeh [12] developed a simple method for prediction of the flash points of unsaturated hydrocarbons including alkenes, alkynes and aromatics. Pan et al. [13] constructed models of the relationships between the structures and flash points of 92 alkanes by means of ANNs using the group bond contribution method. Simulated with the final optimum BP ANN [9-5-1], the results showed that the predicted flash points were in good agreement with the experimental data. Khajeh and Modarress [14] developed two QSPR models to predict the flash points of ester compounds by using genetic function approximation (GFA) and adaptive neuro-fuzzy inference system (ANFIS) methods. The squared correlation coefficients of 0.969 for ANFIS and 0.965 for GFA for testing set showed that these methods had good predictive ability and robustness for estimating flash points of esters. Recently, we developed successfully a QSPR model for the flash points of hydrocarbon based on novel topological

* Corresponding authors. Tel./fax: +86 731 85258733.

E-mail addresses: yimindai@sohu.com (D. Yimin), zhangyuefei74@163.com (Z. Yuefei).

electro-negativity indices Y_C and W_C and path number parameter P_3 [15]. The average absolute error of flash points was only 3.86K between the experimental and the calculated values, and the relative error was 1.46%.

In this study, a new group contribution-based model is presented for the estimation of the flash points of ester compounds.

2. Methodology

2.1. Dataset

Flash point dataset of 110 esters was taken from literatures [16–18]. Flash point values were in the range from 246 to 489K. Initially, the dataset was randomly divided into two groups using Kennard and Stones method [19], namely, one with 80 compounds used for developing the model (training set) and the other with 30 compounds used for validating the obtained model (testing set) and testing its prediction ability. Kennard and Stones method starts by finding two samples, based on the input variables that are the farthest apart from each other. The two samples are removed from the original dataset and put into the calibration set. This procedure is repeated until the desired number of samples has been selected in the calibration set [20,21]. This splitting method can guarantee that the prediction set span the entire range of the experimental measurements and is numerically representative of the dataset.

2.2. Group contribution-based descriptor

A current encouraging trend in predicting the flash point from their chemical structure is the using of empirical correlation [22, 23], topological index [24–26], and the group contribution method [27–29]. One of the most widely useful flash point estimation methods for complex molecules is group contribution method, since this method combined together information on both group property and connectivity in the analyzed molecules. It is based on the assumption that the contribution of a certain group is completely the same in different molecules, and the properties of compounds are considered as the contribution addition of groups which constitute the compounds [30].

In this work, the flash point values of ester compounds were regarded as the concerted contribution of numerous groups constituting the ester molecules. There are six species of groups: primary (CH_3-), secondary ($\text{CH}_2<$), ternary ($-\text{CH}<$), quaternary ($>\text{C}<$) carbons, being symbolic in C_1 , C_2 , C_3 and C_4 , respectively. Another ester compounds include two series groups of ($-\text{O}-$) and ($\text{O}=\text{}$). Contributions for compounds contain CH_3- , $\text{CH}_2<$, $-\text{CH}<$, $>\text{C}<$, $-\text{O}-$ and $\text{O}=\text{}$ six groups. Thus, it is firstly assumed that the flash point values can be formulated as a function of the above groups as follows:

$$\text{FP} = f(\lambda_{\text{CH}_3-}, \lambda_{\text{CH}_2<}, \lambda_{-\text{CH}<}, \lambda_{>\text{C}<}, \lambda_{-\text{O}-}, \lambda_{\text{O}=\text{}}). \quad (1)$$

The distance matrix \mathbf{D} of n atoms in an ester molecular, a square symmetric matrix, can be expressed as $\mathbf{D} = [d_{ij}]_{n \times n}$, where d_{ij} is the length of the shortest path between the vertices i and j in the molecule. We used the relative bond length of two adjacent vertices to distinguish saturated, unsaturated bond and heteroatom compound [31]. Here, d_{ij} is calculated by summing the bond length between two adjacent vertices in the shortest path. If we employ C–C bond length $L_{C-C} = 0.154$ nm as 1, and the relative bond length between vertices i and j is calculated: $d_{ij} = \sum L_{ij}/L_{C-C}$. For example, C=C relative bond length is $0.134/0.154 = 0.8701$. Some kinds of relative bond length are given in Table 1 [32].

Table 1
Several kinds of chemical bonds and their relative bond lengths.

Bond	L/L_{C-C}	Bond	L/L_{C-C}
C=C	0.134/0.154	C ≈ C(benz) ^a	0.139/0.154
C–O	0.143/0.154	C=O	0.122/0.154
C≡C	0.120/0.154		

^a ≈:Conjugation bond.

The distance matrix \mathbf{D} can be expressed as follows:

$$\mathbf{D} = \begin{bmatrix} 0 & L_{12}/L_{C-C} & \cdots & L_{1(n-1)}/L_{C-C} & L_{1n}/L_{C-C} \\ L_{21}/L_{C-C} & 0 & \cdots & L_{2(n-1)}/L_{C-C} & L_{2n}/L_{C-C} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L_{i1}/L_{C-C} & \cdots & 0 & \cdots & L_{in}/L_{C-C} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L_{n1}/L_{C-C} & L_{n2}/L_{C-C} & \cdots & L_{n(n-1)}/L_{C-C} & 0 \end{bmatrix}. \quad (2)$$

Electric charges in the molecule are obviously responsible for electrostatic interactions. As one of the main physicochemical parameters of atoms, electro-negativity represents the ability of atoms to obtain or lose electrons [33]. In order to reflect the chemical information at the atom and group levels, in this work Pauling electro-negativity is defined as electro-negativity matrix. Meanwhile, vertex of atom in a molecule reflects molecular size and the degree of branching. Double bond is regarded as two single bonds and three bonds as three single bonds while calculating [34]. The relative bond length between two adjacent vertices and electro-negativity of atoms is used to present the properties and interaction of the vertices in a molecule. In addition, addition matrix \mathbf{S} , vertex matrix \mathbf{R} and electro-negativity matrix \mathbf{X} are defined in order to distinguish the length of carbon backbone, the size, the degree of branching and the charge distribution of a molecule, respectively [35]. Where s , r , and x are determined by the sum of the distance of relative bond length of the shortest path between the vertices, the degree of branching in corresponding vertices and the Pauling electro-negativity of atom, respectively. According to the corresponding definition matrix \mathbf{S} , \mathbf{R} and \mathbf{X} are expressed as follows:

$$\mathbf{S} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}, \quad s_i = \sum_{j=1}^n d_{ij}, \quad \mathbf{R} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_n \end{bmatrix}. \quad (3)$$

The matrix of any group i is represented by an n -dimensional vector matrix to characterize the position of the group in the molecule. The module λ_i for any point vector is defined as follows:

$$\lambda_i = \log(S_i \cdot r_i \cdot \chi_i). \quad (4)$$

For special series group in molecule, new group contribution-based descriptor is defined as follows:

$$\sum \lambda_i = \log \sum (S_i \cdot r_i \cdot \chi_i). \quad (5)$$

For saturated esters, the contributions of the properties arising from six parts, namely, the contributions of CH_3- , $\text{CH}_2<$, $-\text{CH}<$, $>\text{C}<$, $-\text{O}-$ and $\text{O}=\text{}$, are respectively expressed as:

$$\kappa_5 = \sum \lambda_{-\text{O}-}, \kappa_6 = \sum \lambda_{\text{O}=\text{}}, \\ \kappa_1 = \sum \lambda_{\text{CH}_3-}, \kappa_2 = \sum \lambda_{\text{CH}_2<}, \kappa_3 = \sum \lambda_{-\text{CH}<}, \kappa_4 = \sum \lambda_{>\text{C}<}, \kappa_5 = \sum \lambda_{-\text{O}-}, \kappa_6 = \sum \lambda_{\text{O}=\text{}}.$$

For unsaturated esters, if double bonds and three bonds are located in the middle of molecules, bonds are treated as one chemical bond while calculating. While double bonds and three bonds are the ends of

Download English Version:

<https://daneshyari.com/en/article/1179786>

Download Persian Version:

<https://daneshyari.com/article/1179786>

[Daneshyari.com](https://daneshyari.com)