Process Safety and Environmental Protection xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Process Safety and Environmental Protection

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

# MNLR and ANN structural group contribution methods for predicting the flash point temperature of pure compounds in the transportation fuels range

*Tareq A. Albahri* *

*Chemical Engineering Department, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait*

ABSTRACT

A QSPR method is presented for predicting the flash point temperature (FPT) of pure compounds in the transportation fuels range. A structural group contribution method is used to determine the flash point temperature using two techniques: multivariable nonlinear regression and artificial neural networks. The method was used to probe the structural groups that have significant contribution to the overall FPT of pure compounds and arrive at the set of 37 atom-type structural groups that can best represent the flash point for about 375 substances. The input parameters to the model are the number of occurrence of each of the 37 structural groups in each molecule. The neural network method was the better of the two techniques and can predict the flash point of pure compounds merely from the knowledge of the molecular structure with an overall correlation coefficient of 0.996 and overall average and maximum errors of 1.12% and 6.62%, respectively. The results are compared to the more traditional approach of the SGC method along with other methods in the literature.



*Keywords:* Flash point; Group contribution; Molecular modeling; Neural networks; Quantitative structure property relation; QSPR

## 1. Introduction

In recent years, the term environmental impact has extended its traditional meaning to include other extensive concepts in view of the possibility of industrial accidents which, because of their magnitude, are capable of causing significant damage to people and the environment. This concern, which in the past was principally associated with the nuclear industry, now includes the chemical industry and their safety. Among these concerns are incidents of fire disasters in chemical and petrochemical plants caused by the leak of materials at or above their auto ignition temperature or flash point or within their flammability limits.

The flammability characteristics of chemical substances are very important for safety considerations in storage, processing, and handling. These characteristics which include the flash point, the auto ignition temperature, and the upper and lower flammability limits are some of the most important safety specifications that must be considered in assessing the overall flammability hazard potential of a chemical substance, defined as the degree of susceptibility to ignition or release of energy under varying environmental conditions. Experimental values of these properties are always desirable, however, they are scarce and expensive to obtain. When experimental values are not available and determining them by experimental means is not practical, a prediction method, which is desirably convenient and fast, must be used to estimate them.

The flash point temperature (FPT) is one of the most important safety specifications used to characterize the hazard potential of a chemical substance. The flash point

* Tel.: +965 2481 7662; fax: +965 2483 9498.
E-mail address: toalbahri@gmail.com

Received 1 December 2013; Received in revised form 27 January 2014; Accepted 15 March 2014
http://dx.doi.org/10.1016/j.psep.2014.03.005
0957-5820/

temperature of a combustible substance is the minimum temperature at which sufficient vapor is produced to form an ignitable mixture with air near the surface of the liquid or within the vessel used which is within the flammability limits. This is not to be confused with the fire point where the vapor will continue to burn even after the removal of the ignition source; at the flash point, once the ignition source is removed, the fire will stop.

FPT can indicate the possible presence of highly volatile or flammable materials in a relatively nonvolatile (nonflammable) environment and is usually used in shipping and safety regulations to define and classify flammable and combustible materials. The U.S. Department of Transportation (DOT) and the U.S. Department of Labor (OSHA) have established that liquids with a flash point under 37.8 °C (100 °F) are considered flammable.

There exist several recognized tests for evaluating the PFT of a substance or fuel, which differ according to the characteristics of the liquid under study. Standard ASTM (2002) closed-cup test methods include Tag (D56-01), small scale (D3828-98), Setaflash (D3828), Pensky-Martens (D93-00), and the equilibrium method (D3941-90). Standard ASTM (2002) open-cup test methods include Cleveland (D92-01) and Tag (D1310). Generally, methods using closed vessels are used for low FPT substances and give less value than those obtained with open vessels, although the differences are small. In all cases the procedure involves a slow heating of the liquid in contact with the air, applying a source of ignition at predetermined intervals and recording the temperature at which burning occurs.

### 1.1. *Background*

When the flammability characteristics cannot be determined experimentally, empirical equations for their determination are available. Detailed review of the FPT predictions methods has been discussed extensively in the literature. Gharagheizi et al. (2011) for example, developed a simple empirical correlation to predict FPT using the normal boiling point and the number of carbon atoms. Although the above correlation was able to predict FPT for 1471 pure organic compounds with an absolute average deviation of 2.4% and a correlation coefficient of 0.979, it requires the normal boiling point to predict the flash point which is not always available or convenient.

Albahri (2003a) introduced the concept of using structural group contributions (SGC) to predict FPT, from the molecular structure of the compound alone, using multivariable nonlinear regression (MNLR). The input parameters to the final polynomial equation are the number of occurrence of each of the structural groups in the molecule in addition to the group contribution values. Although the model was able to accurately predicted FPT for 287 compounds with a correlation coefficient of 0.98 and average error of 1.2%, it is limited to hydrocarbons only.

Keshavarz and Ghanbarzadeh (2011) developed a simple method for predicting the flash point of unsaturated hydrocarbons including alkenes, alkynes and aromatics with an absolute average deviation of 11 K. The number of carbon and hydrogen atoms is used as a core function that can be revised for some compounds using a structural parameter correcting function. The method was developed using a set of 173 compounds and tested using another set of 76 compounds and is restricted to unsaturated hydrocarbons alone which limits its applicability.

Rowley et al. (2010) developed a correlative method for estimating the flash point of 1062 organic compounds based entirely on structural contributions. The proposed correlation based on Clausius–Clapeyron equation and Leslie–Geniessee relation results in an average absolute error and deviation of 2.84% and 9.8 K, respectively. The sum of the 62 atomic structural group contributions is used to predict the boiling point input parameter of the developed correlation. Although the method is accurate enough, the structural group definitions and restrictions make the method hard to practice. Rowley et al. (2011) developed yet another correlation for predicting the flash point of the same set of pure organic compounds based on the normal boiling point and enthalpy of vaporization at the normal boiling point with an absolute average error and deviation of 1.32% and 4.65 K, respectively. The method requires knowledge of other thermo-physical properties which are inconvenient. Furthermore, significant errors are reported for carboxylic acids.

Li and Liu (2010) used Le Chatelier's rule and Antoine equation to correlate the vapor pressure and flash point, and further provided a comprehensive review of the flash point prediction methods based on vapor pressure, molecular structure, composition range, and boiling point. They ultimately recommended using QSPR with artificial neural networks as a correlation technique because of its nonlinear property, high accuracy, and potential for wide application.

In terms of using artificial intelligence, Saldana et al. (2011) developed a method to predicting the flash point of 437 hydrocarbons, alcohols and esters using QSPR methods. Various approaches were investigated from linear modeling such as genetic function approximation (GA) and partial least squares (PLS) to nonlinear methods such as feed-forward artificial neural networks (FF-ANN), general regression neural networks (GRNN), support vector mechanics (SVM) and graph machines (GM). None of the models was significantly more accurate than the others; thus, consensus modeling was used to improve generalization and predictive power compared to individual predictive models. The correlation coefficient was 0.922 and the absolute average error and deviation were 3.2% and 10.4 K, respectively. The consensus method is the average of all the above predictive methods which make the method cumbersome for practice.

Khajeh and Modarress (2011) developed a QSPR model to predict the flash point of 151 alcohols using genetic function approximation (GFA) and using four molecular descriptors as input to adaptive neuro-fuzzy inference system (ANFIS) model with a correlation coefficient of 0.931 and 0.957, respectively. However, the method is limited to only alcohols.

Gharagheizi and Alamdari (2008) used a QSPR model and Genetic Algorithm-based Multivariate Linear Regression (GA-MLR) technique to select four chemical structure-based molecular descriptors to predict FPT of 1030 organic compounds with a correlation coefficient of 0.9669 and average deviation of 12.691 K. Gharagheizi et al. (2008) further used 79 structural groups in a 79-9-1 feed forward neural network to correlate FPT of 1378 organic compounds with a correlation coefficient of 0.9757, and an average absolute deviation and maximum error of 8.1 K and 26%, respectively. The four molecular descriptors and the 79 structural groups are both intricate and hard to determine for each molecule which poses difficulties on the methods applicability in practice.

Pan et al. (2007) developed a back-propagation 9-5-1 neural network model using group-bond contribution method to model FPT for 92 alkanes. Although the model was accurate