Journal of Hazardous Materials 170 (2009) 595–604

Contents lists available at ScienceDirect

Journal of Hazardous Materials

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

A new group contribution-based model for estimation of lower flammability limit of pure compounds

Farhad Gharagheizi [a,b,*]

[a] Department of Chemical Engineering, Faculty of Engineering, University of Tehran, P.O. Box 11365-4563, Tehran, Iran
[b] Department of Chemical Engineering, Medicinal Plants and Drugs Research Institute, Shahid Beheshti University, Evin, Tehran, Iran

ARTICLE INFO

Article history:
Received 21 January 2009
Received in revised form 25 April 2009
Accepted 4 May 2009
Available online 15 May 2009

Keywords:
Lower flammability limit
Group contributions
Neural network
Safety operational condition

ABSTRACT

In the present study, a new method is presented for estimation of lower flammability limit (LFL) of pure compounds. This method is based on a combination of a group contribution method and neural networks. The parameters of the model are the occurrences of a new collection of 105 functional groups. Basing on these 105 functional groups, a feed forward neural network is presented to estimate the LFL of pure compounds. The average absolute deviation error obtained over 1057 pure compounds is 4.62%. Therefore, the model is an accurate model and can be used to predict the LFL of a wide range of pure compounds.

1. Introduction

Flammability characteristics of chemical compounds are needed to design safe operational conditions in the chemical and petrochemical plants [1,2]. One of the most important flammability characteristics is lower flammability limit (LFL) of pure compounds in air. Every combustible gas burns in air only over a limited range of concentration. Lower than an especial concentration of the compound in air which, is called lower flammability limit, the mixture of the compound with air is too lean, and while above another especial concentration which, is called upper flammability limit (UFL) the mixture is too rich. The concentrations between these two limits constitute the flammable range. Therefore, to prevent from fire and explosion of a flammable gas, knowledge about LFL is critical.

The LFL depends on several factors such as nature of the compound, the geometry of the apparatus, strength of the ignition source, the test temperature and pressure, degree of mixing, oxygen concentration, and concentration of the diluents [3–5]. Therefore measuring the LFL requires the standard apparatus and several conditions as stated in ASTM-E681. On the other hands, the reported values of the LFL in the literature differs each other because they are not measured in the exact conditions of ASTM-E681.

According to the ASTM-E681, measuring the LFL is time-consuming and expensive; therefore, application of computational methods is necessary to develop an accurate method for estimation of the property.

Several methods have been presented for estimation and prediction of the LFL of pure compounds. Spakowski presented a model for estimation of LFL based on standard heat of combustion (ΔH_{comb}) [1,6]. The model is:

$$\text{LFL}\,(\%\text{vol}) = -\frac{4354}{\Delta H_{comb}}(\text{kJ/mol}) \quad (1)$$

As reported by Albahri [7], application of this method for estimation of 454 pure compounds respectively shows average deviation, maximum deviation, average error, and squared correlation coefficient of 1.35 (vol%), 14.02 (vol%), 12.3 (vol%), and 0.83. Jones presented another method for estimation of LFL of pure compounds based on the concentration of the flammable product for complete combustion in air (C_{est}). This model is shown in Eq. (2).

$$\text{LFL}\,(\%\text{vol}) = 0.55C_{est} \quad (2)$$

This model showed better results in comparison with the Spakowski's method presented in Eq. (1). Base on evaluations of Albahri [7], the Jones' method respectively shows average deviation, maximum deviation, average error, and squared correlation coefficient of 0.07 (vol%), 5.7 (vol%), 6.13 (vol%), and 0.89 over the same 454 pure compounds used to evaluate Spakowski's method.

As stated by Sheldon [4], these two methods are only approximate and fail with low molecular weight compounds. Albahri [7] presented a structural group contribution method for estimation of LFL of pure compounds. In this model, 19 simple functional groups were used to develop a model for estimation of LFL of pure

* Tel: +98 21 66957784; fax: +98 21 66957784.
E-mail addresses: fghara@ut.ac.ir, fghara@gmail.com.

0304-3894/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.jhazmat.2009.05.023

compounds. This model respectively shows average deviation, maximum deviation, average error, and squared correlation coefficient of 0.04 (vol%), 5.6 (vol%), 4.1 (vol%), and 0.93 over the same 454 pure compounds used to evaluate two previous models (Spakowski's method and Jones' methods).

Recently, a quantitative structure-property relationship was presented by one of the authors for prediction of LFL of pure compounds [8]. To develop this model, 1057 pure compounds were used. The obtained model respectively shows average deviation, maximum deviation, average error, and squared correlation coefficient of 0.35 (vol%), 3.36 (vol%), 7.8 (vol%), and 0.97 over all 1057 pure compounds.

All these methods are useful, but they have some disadvantages. Spakowski's method and Jones' method are very approximate and obtained based on a small group of compounds. These types of methods cannot generally be used for estimation of LFL. Perhaps, the method presented by Albahri is the first accurate method for estimation of LFL of pure compounds but, this method presented only for hydrocarbons therefore, application of this method for other compounds are not possible. The presented method by the author is a comprehensive method (comprehensive means that this method has no basic limitation in use for the chemical families of compounds) but, the method is not easy to use because the complex procedure for computation of parameters.

The aim of this study is to present a model based on a combination of a new collection of group contributions (for description of molecular structure of pure compounds) and neural networks (to obtain an accurate model) for estimation of LFL of pure compounds. Perhaps, group contribution methods are simplest methods which use only chemical structure of compounds for estimation of variety of properties in science and engineering. Therefore, major aim in this study is to present simpler and more accurate method than previously presented methods for estimation of LFL of pure compounds.

2. Dataset preparation

The quality of an estimation method directly depends on the quality of the dataset used for its development. There are many compilations for physical properties of pure compounds but, of them, DIPPR 801 [9] has some advantages. This database is the result of a vast literature survey performed under supervision of American Institute of Chemical Engineers (AIChE). The most important advantage of this database is the performed evaluations over all collected values. The result of these evaluations is the unique recommended values for every physical property. Application of the database for this study is found very useful therefore, 1057 pure compounds were found in this database and used for this study. This dataset is the same like as the dataset used by the author in previously presented model for prediction of LFL of pure compounds.

2.1. Development of new group contributions

In this step, the chemical structures of all 1057 compounds were analyzed and finally, 105 functionally groups were found useful to estimate the LFL. Perhaps, these functional groups are simplest functional groups selected from those functional groups proposed and used by various researchers in various versions of group contribution methods for various physical properties. Application of these functional groups showed promising results in prediction of previous flammability properties of pure compounds. Therefore these parameters are used to present a new model for prediction of LFL of pure compounds.

The functional groups found and used in this study and their chemical structures are extensively presented in Table 1.

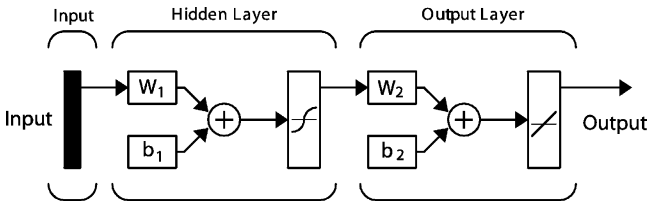

Fig. 1. The schematics structure of the three-layer feed forward neural network used in this study.

These 105 functional groups and their numbers of occurrences in pure compounds are presented as supplementary materials. These functional groups are used as input parameters for the model.

2.2. Generation of neural network based-group contribution

When the group contributions table was provided, we should find a correlation between these groups and the LFL of pure compounds. The simplest method is to assume multi-linear relationship between these groups and the LFL. This solution is the same method, used in the classic group contribution technique. Application of this methodology for this problem is failed. We could not find a good model by this method. Therefore, application of nonlinear methods such as neural networks was considered useful for this problem.

Neural networks are extensively used in various scientific and engineering areas such as estimations of physical and chemical properties [10]. These powerful tools are usually applied to study of the complicated systems such as the problem defined here. The theoretical explanations about neural networks can be found in many references such as ref. [11].

This solution is found useful and therefore, using the Neural Network toolbox of the MATLAB software (Mathworks Inc. software), three layer feed forward neural networks were evaluated for the problem. The schematic typical structure of three layer feed forward neural networks is presented in Fig. 1.

This type of neural networks has been used by one of the authors in his previous works, therefore, the detail explanations about the three layer feed forward used in this study can be found, elsewhere [12–19]. The simplified form of the relationship between input parameters and output of a three-layer FFNN can be shown as Eq. (3).

$$y_{\mathrm{calc}}(i) = (W_2 \times (\tanh((W_1 \times T_{\mathrm{i}}) + b_1)) + b_2 \quad (3)$$

In this equation, T is the input matrix of dimension $nparam \times nds$. $nparam$ is the number of functional groups (it is equal 105 in this study) and nds is the number of available compounds of the training set (it is equal 846 in this study). T_{i} is the ith-column of the Matrix T. W_1 is the fist weight matrix of the three layer FFNN and is of dimension $n \times nds$. n is the number of neurons in the hidden layer. b_1 is the first bias matrix of dimension $n \times 1$. W_2 is the second weight matrix of output layer and is of dimension $n \times 1$. b_2 is the second bias of output layer which is a scalar value. $y_{\mathrm{calc}}(i)$ is the ith-output of this network which should be compared with ith-member of the property.

All the 105 functional groups and the LFL values should be normalized between −1 and +1 to decrease computational errors. This work can be performed using maximum and minimum values of every 105 functional groups for inputs and using maximum and minimum values of the LFL for output. After this step, the main dataset should be divided into two new datasets. These two datasets include: training set and test set. The training set used to generate and optimized neural networks and the test set is used only to check validity of the obtained model. The process of division of main dataset into two new datasets is usually randomly performed. For this purpose, 80% of the main dataset randomly selected for the