



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.

© 2009 Published by Elsevier B.V.

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:

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

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).

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

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: fggara@ut.ac.ir, fggara@gmail.com.

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.

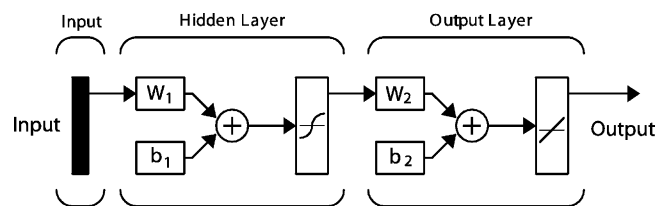


Fig. 1. The schematics structure of the three-layer feed forward neural network used in this study.

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).

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

In this equation, T is the input matrix of dimension $n_{\text{param}} \times n_{\text{ds}}$. n_{param} is the number of functional groups (it is equal 105 in this study) and n_{ds} is the number of available compounds of the training set (it is equal 846 in this study). T_i is the i th-column of the Matrix T . W_1 is the first weight matrix of the three layer FFNN and is of dimension $n \times n_{\text{ds}}$. 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. $y_{\text{calc}}(i)$ is the i th-output of this network which should be compared with i th-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

Table 1
Functional groups used to develop the model.

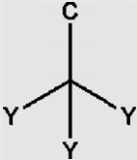
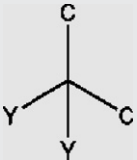
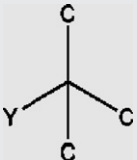
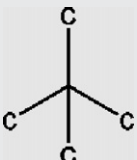
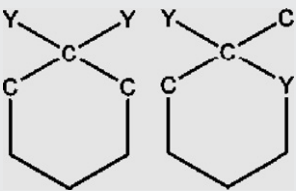
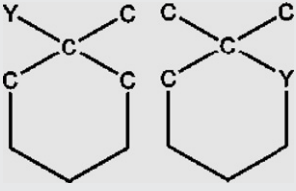
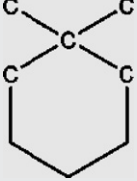
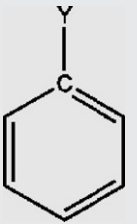
No.	ID	Functional groups	Comments
1	LFL001		Number of terminal primary C(sp3) Y = any terminal atom or heteroaromatic group (i.e. H, X, OH, NH2, etc.)
2	LFL002		Number of total secondary C(sp3) Y = H or any heteroatom
3	LFL003		Number of total tertiary C(sp3) Y = H or any heteroatom
4	LFL004		Number of total quaternary C(sp3)
5	LFL005		Number of ring secondary C(sp3) Y = H or any heteroatom
6	LFL006		number of ring tertiary C(sp3) Y = H or any heteroatom
7	LFL007		Number of ring quaternary C(sp3)
8	LFL008	Sum of all the carbons belonging to any aromatic and heteroaromatic structure	Number of aromatic C(sp2)
9	LFL009		Number of substituted benzene C(sp2) Y = carbon or any heteroatom

Table 1 (Continued)

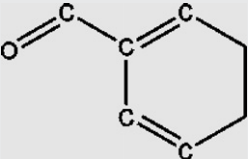
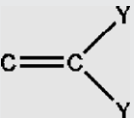
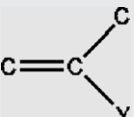
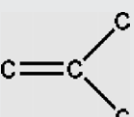
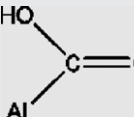
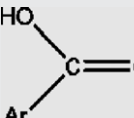
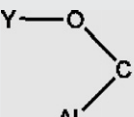
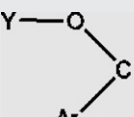
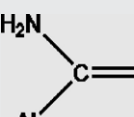
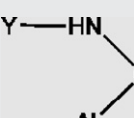
No.	ID	Functional groups	Comments
10	LFL010		Number of non-aromatic conjugated C(sp ²)
11	LFL011		Number of terminal primary C(sp ²) Y = any terminal atom or heteroaromatic group (i.e. H, X, OH, NH ₂ , etc.)
12	LFL012		number of aliphatic secondary C(sp ²) Y = H or any heteroatom
13	LFL013		Number of aliphatic tertiary C(sp ²)
14	LFL014	$C=C=C$	Number of allenes groups
15	LFL015	$Y-C\equiv C$	Number of terminal C(sp) Y = any terminal atom or heteroaromatic group (i.e. H, X, OH, NH ₂ , etc.)
16	LFL016	$Y-C\equiv C$	Number of non-terminal C(sp) Y = C or any non-terminal heteroatom
17	LFL017	$Al-N=C=O$	Number of isocyanates (aliphatic)
18	LFL018		Number of carboxylic acids (aliphatic)
19	LFL019		Number of carboxylic acids (aromatic)
20	LFL020		Number of esters (aliphatic) Y = Ar or Al (not H) Al = H or aliphatic group linked through C
21	LFL021		Number of esters (aromatic) Y = Al or Ar
22	LFL022		Number of primary amides (aliphatic) Al = H or aliphatic group linked through C
23	LFL023		Number of secondary amides (aliphatic) Y = Ar or Al (not H, not C=O) Al = H or aliphatic group linked through C

Table 1 (Continued)

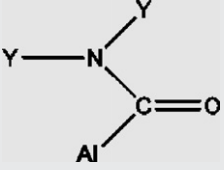
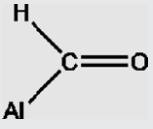
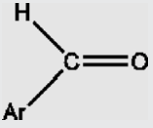
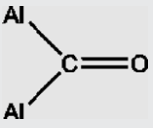
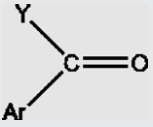
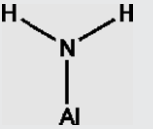
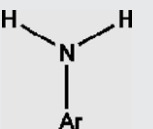
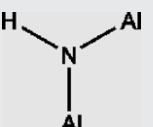
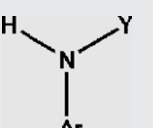
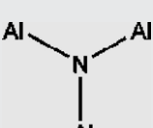
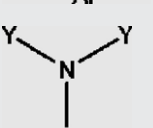
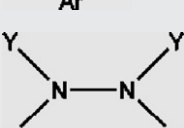
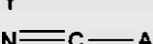
No.	ID	Functional groups	Comments
24	LFL024		Number of tertiary amides (aliphatic) Y = Ar or Al (not H, not C=O) Al = H or aliphatic group linked through C
25	LFL025		Number of aldehydes (aliphatic)
26	LFL026		Number of aldehydes (aromatic)
27	LFL027		Number of ketones (aliphatic)
28	LFL028		Number of ketones (aromatic) Y = Al or Ar
29	LFL029		Number of primary amines (aliphatic) Al = aliphatic group linked through C (not C=O)
30	LFL030		Number of primary amines (aromatic)
31	LFL031		Number of secondary amines (aliphatic) Al = aliphatic group linked through C (not C=O)
32	LFL032		Number of secondary amines (aromatic) Y = Ar or Al (not C=O)
33	LFL033		Number of tertiary amines (aliphatic) Al = aliphatic group linked through C (not C=O)
34	LFL034		Number of tertiary amines (aromatic) Y = Ar or Al (not C=O)
35	LFL035		Number of N hydrazines Y = C or H
36	LFL036		Number of nitriles (aliphatic)

Table 1 (Continued)

No.	ID	Functional groups	Comments
37	LFL037		Number of positive charged N
38	LFL038	$\text{Al}-\text{O}-\text{H}$	Number of hydroxyl groups Al = aliphatic group linked through any atom
39	LFL039	$\text{Ar}-\text{O}-\text{H}$	Number of aromatic hydroxyls Ar = aromatic group linked through any atom
40	LFL040		Number of primary alcohols
41	LFL041		Number of secondary alcohols
42	LFL042		Number of tertiary alcohols
43	LFL043	$\text{Al}-\text{O}-\text{Al}$	Number of ethers (aliphatic) Al = aliphatic group linked through C (not C=O, not C # N)
44	LFL044	$\text{Ar}-\text{O}-\text{Y}$	Number of ethers (aromatic) Y = Ar or Al (not C=O, not C # N)
45	LFL045		Number of anhydrides (thio-) Y = O or S
46	LFL046		Number of thiols
47	LFL047	$\text{C}-\text{S}-\text{C}$	Number of sulfides
48	LFL048	$\text{C}-\text{S}-\text{S}-\text{C}$	Number of disulfides
49	LFL049		Number of sulfates (thio-/dithio-) Y = O or S
50	LFL050		

Table 1 (Continued)

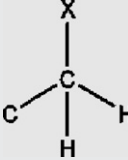
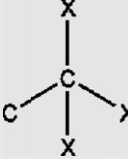

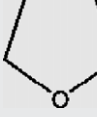
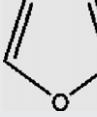
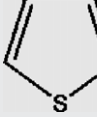
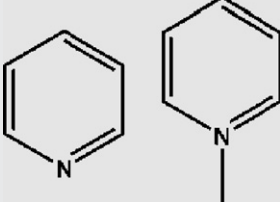
No.	ID	Functional groups	Comments
51	LFL051		Number of CH2RX
52	LFL052		Number of CRX3
53	LFL053	Ar—X	Number of X on aromatic ring
54	LFL054		Number of oxiranes
55	LFL055		Number of oxolanes
56	LFL056		Number of furanes
57	LFL057		Number of thiophenes
58	LFL058		Number of pyridines
59	LFL059	Sum of the hydrogens linked to all of the Os and Ns in the molecule	Number of donor atoms for H-bonds (N and O)
60	LFL060	Total number of Ns, Os and Fs in the molecule, excluding N with a formal positive charge, higher oxidation states and pyrrolyl form of N	Number of acceptor atoms for H-bonds (N, O, F)
61	LFL061	CH3R/CH4	
62	LFL062	CHR3	
63	LFL063	CH3X	
64	LFL064	CH2RX	
65	LFL065	CHR2X	
66	LFL066	CHRX2	
67	LFL067	CR3X	
68	LFL068	=CHX	
69	LFL069	=CRX	
70	LFL070	#CR/R=C=R	
71	LFL071	R···CR···R	
72	LFL072	R···CX···R	
73	LFL073	R···CH···X	
74	LFL074	R···CR···X	
75	LFL075	R···CH···X	
76	LFL076	R···CR···X	
77	LFL077	Al-CH=X	
78	LFL078	R-C(=X)-X/R-C#X/X=C=X	
79	LFL079	H ^a attached to C ⁰ (sp ³) no X attached to next C	
80	LFL080	H ^a attached to C ¹ (sp ³)/C ⁰ (sp ²)	
81	LFL081	H ^a attached to C ² (sp ³)/C ¹ (sp ²)/C ⁰ (sp)	
82	LFL082	H ^a attached to C ³ (sp ³)/C ² (sp ²)/C ³ (sp ²)/C ³ (sp)	

Table 1 (Continued)

No.	ID	Functional groups	Comments
83	LFL083	H attached to alpha-C ^b	
84	LFL084	H ^a attached to C ⁰ (sp ³) with 1X attached to next C	
85	LFL085	H ^a attached to C ⁰ (sp ³) with 2X attached to next C	
86	LFL086	Alcohol	
87	LFL087	Phenol/enol/carboxyl OH	
88	LFL088	=O	
89	LFL089	Al–O–Al	
90	LFL090	Al–O–Ar/Ar–O–Ar/R [·] ·O [·] ·R/R–O–C=X	
91	LFL091	R–O–O–R	
92	LFL092	Ar–NH–Al	
93	LFL093	Ar–NAI ₂	
94	LFL094	RCO–N</>N–X=X	
95	LFL095	Ar ₂ NH/Ar ₃ N/Ar ₂ N–Al/R [·] ·N [·] ·R ^c	
96	LFL096	R#N/R=N–	
97	LFL097	R [·] ·N [·] ·R/R [·] ·N [·] ·X ^c	
98	LFL098	F ^a attached to C ³ (sp ³)	
99	LFL099	Cl ^a attached to C ¹ (sp ³)	
100	LFL100	Cl ^a attached to C ¹ (sp ²)	
101	LFL101	Cl ^a attached to C ² (sp ²)-C ⁴ (sp ²)/C ¹ (sp)/C ⁴ (sp ³)/X	
102	LFL102	Br ^a attached to C ¹ (sp ³)	
103	LFL103	R ₂ S/RS–SR	
104	LFL104	R–SO–R	
105	LFL105	R–SO ₂ –R	

Explanations: R represents any group linked through carbon; X represents any electronegative atom (O, N, S, P, Se, halogens); Al and Ar represent aliphatic and aromatic groups, respectively; = represents a double bond; # represents a triple bond; ··· represents an aromatic bond as in benzene or delocalized bonds such as the N–O bond in a nitro group; ··· represents aromatic single bonds as the C–N bond in pyrrole.

^a The superscript represents the formal oxidation number. The formal oxidation number of a carbon atom equals the sum of the conventional bond orders with electronegative atoms; the C[·]·N bond order in pyridine may be considered as 2 while we have one such bond and 1.5 when we have two such bonds; the C[·]·X bond order in pyrrole or furan may be considered as 1.

^b An alpha-C may be defined as a C attached through a single bond with –C=X, –C#X, –C–X.

^c Pyrrole-type structure.

training set and all other 20% is used for the test set. Regarding the allocation percent of the test set from main dataset, it should be noted that the researchers have been used various allocation percents of main dataset for the test set. For example, Albahri and George [20] used 5% of the main dataset for the test set (20 compounds from 490 compounds). In another work, Albahri [21] used 5% of the main dataset for the test set (9 compounds from 200 compounds). Besides these works, Gharagheizi et al. [8] used 10% of the main dataset for the test set (137 compounds from 1378 compounds). The effect of the allocation percent of test set from the main dataset on the accuracy of the neural networks has been studied by the author [22]. The results of this study show that the percent of test set allocated from the main dataset should be between 5–35%. If this percent is lower than 5% the accuracy of the model over the training set is greater than the test set. Also, if the percent is greater than 40% the obtained model cannot predict the test set as well as the training set. On the other hand the experiences of the author show that the optimum percent of the test set is dependent to the nature of the problem. During solving a problem it should be used different percents of main dataset for the test set. The optimum percent is the percent which the accuracy of the model over the test set approaches the training set.

Various percents of the main dataset were used for the test set. Of them, for several times, application of 20% of the main dataset for the test set showed the least difference between accuracies of the model over the test set and the training set. As a result, 20% of the main dataset was used to the test set and all the others were used to the training set (211 compounds for the test set and 846 compounds for the training set). These compounds were randomly selected.

Using the train set and the test set the three layer feed forward neural networks were applied to generate a model to predict LFL.

Generating a neural network means determination of the weight matrices and bias vectors. As shown in the Fig. 1, there are two weight matrices and two bias vectors in a three layer feed forward neural network; W_1 and W_2 , b_1 and b_2 . These parameters should

be obtained by minimization of an objective function. The objective function used in this study is the sum of squares of error between the outputs of the neural network (estimated LFL) and the target values (real LFL of those compounds). This minimization was performed by Levenberg–Marquardt algorithm. This algorithm is rapid and accurate in the process of training neural networks [9,10].

3. Results and discussion

By presented procedure in the previous section, an optimized feed forward neural network was obtained for prediction of LFL. For determination of the number of neurons of hidden layer of the neural network, numbers 1–50 were checked and then the number 5 showed best results. Therefore, the best three layer feed forward neural network has the structure 105–5–1. The mat file (MATLAB file format) of the obtained neural network containing all parameters of the obtained model can be freely accessible by email from the author of this manuscript. The predicted LFL using this model in comparison with the real values are shown in Fig. 2. Also these values are reported as supplementary materials.

As shown in Fig. 2, the accuracy of the predicted LFL over the training set and the test set is almost equal. This case shows that the model is valid. In addition, the predicted LFL of 1057 pure compounds and the DIPPR 801 values are very close to each others. The case shows that the model is accurate in prediction of LFL of pure compounds.

The results obtained by model are presented in Table 2. These results show that the squared correlation coefficient, average absolute deviation, standard deviation error, and root mean square error of the model over the training set and the test set and the main dataset are respectively 0.99%, 0.971%, 0.986%, 4.35%, 5.70%, 4.62%, 0.05%, 0.083%, 0.058%, 0.05%, 0.084%, and 0.058%. The average absolute deviation error obtained by the model over all 1057 compounds is shown in Fig. 3. As can be found, the obtained model is an accurate model to predict the LFL of pure components.

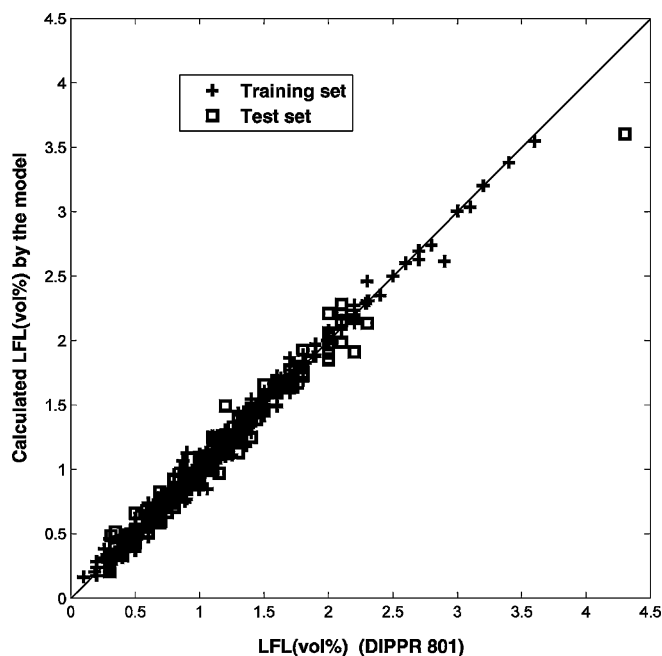


Fig. 2. Comparison between LFL predicted by the model and the DIPPR 801 data.

Table 2

Statistical parameters of the obtained model.

Statistical parameter	Value
Training set	
R^2	0.99
Average absolute deviation	4.35%
Standard deviation error	0.05
Root mean square error	0.05
n	846
Test set	
R^2	0.971
Average absolute deviation	5.70%
Standard deviation error	0.083
Root mean square error	0.084
n	211
Training set + test set	
R^2	0.986
Average absolute deviation	4.62%
Standard deviation error	0.058
Root mean square error	0.058
n	1057

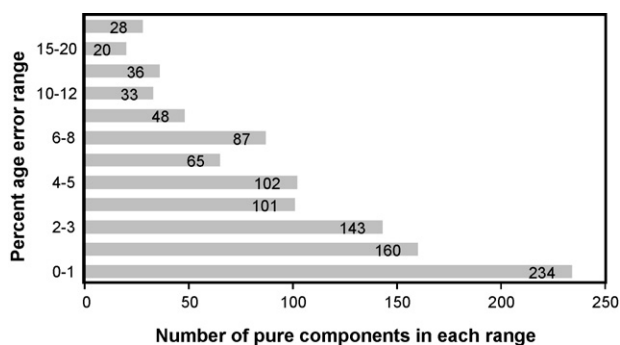


Fig. 3. The percent errors obtained using the presented model and number of pure compounds in each range.

Comparison between the obtained model and previously presented models shows that this model is more comprehensive than Spakowski's method and Jones' method [1,5,6], because it has been validated using a larger dataset. Also, it is more accurate than those. In addition, the obtained model can predict the LFL of low molecular weight compounds that Spakowski's method and Jones' method does not have this capability [4].

In comparison with the Albahri's method [6], this model is more comprehensive because on the contrary of Albahri's method, the applicability range of this model is not limited to the hydrocarbon compounds. Also, this model is validated using a dataset more than two times larger than that used by Albahri (1057 vs. 454 pure compounds). Also, this model shows better squared correlation coefficient.

Also a comparison between this model and previously presented model by one of the authors of this paper [7], shows that this model is more accurate. Also, the computation of parameters of this model is simpler than that model. Therefore, this model estimates LFL of pure compounds more accurate than before.

4. Conclusion

In the presented study, a molecular-based model was presented for prediction of LFL of pure compounds. The model is the result of a combination between group contributions and feed forward neural networks. The needed parameters of the model are the number of occurrences of 105 functional groups in every molecule. It should be noted that many of these 105 functional groups are not simultaneously available in a molecule therefore computation of these parameters from chemical structure of every molecule is simple. For developing the model, 1057 pure compounds were used. Therefore, this model predicts LFL of every regular compound with some limitations. These 1057 pure compounds cover many families of compounds therefore the model has a wide range of applicability but, application of the model is restricted to those compounds similar to the compounds used to develop this model. Application of the model for those compounds which is completely different from compounds used to develop the model is not recommended.

Also, comparison between the presented model with the previously presented models shows that the model in comparison with Spakowski's method and Jones methods and also Albahri's method is more comprehensive and more accurate. Also, the comparison between the model and the QSPR method previously presented by the author (with the same dataset used in this study) shows that this model is more accurate and is simpler to use. Also, this model has lower outliers than that model.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhazmat.2009.05.023.

References

- [1] D.A. Crowl, J.F. Louvar, Chemical Process Safety Fundamentals with Applications, 2nd ed., Prentice-Hall, New Jersey, 2002.
- [2] F.P. Lees, Loss Prevention in the Process Industries, 2, 2nd ed., Butterworth Heinemann, Oxford, UK, 1996.
- [3] Y.N. Shebeko, A.V. Ivanov, E.N. Alekhina, A.A. Barmakova, Calculation of upper limits of concentration for organic compounds clouds inflammation in air, and also heteroatomic compounds, Sov. Chem. Ind. 15 (1983) 599–600.
- [4] M. Sheldon, A Study of the flammability limits of gases and vapors, Fire Prevention 174 (1984) 23–31.
- [5] S.Y. Liao, D.M. Jiang, Z.H. Huang, Q. Cheng, J. Gao, Y. Hu, Approximation of flammability region for natural gas-air-diluent mixtures, J. Hazard. Mater. A125 (2005) 23–38.
- [6] F.P. Bodhurtha, Industrial Explosion Prevention and Protection, McGraw-Hill, New York, 1980.

- [7] T. Albahri, Flammability characteristics of pure hydrocarbons, *Chem. Eng. Sci.* 58 (2003) 3629–3641.
- [8] F. Gharagheizi, Quantitative structure-property relationship for prediction of lower flammability of pure compounds, *Energy Fuels* 22 (2008) 3037–3039.
- [9] Project 801, Evaluated Process Design Data, Public Release Documentation, Design Institute for Physical Properties (DIPPR), American Institute of Chemical Engineers (AIChE) 2006.
- [10] J. Taskinen, J. Yliruusi, Prediction of physicochemical properties based on neural network modelling, *Adv. Drug Deliv. Rev.* 55 (2003) 1163–1183.
- [11] M. Hagan, M.H.B. Demuth, M.H. Beale, *Neural Network Design*, International Thomson Publishing, 2002.
- [12] F. Gharagheizi, R.F. Alamdari, M.T. Angaji, A new neural network-group contribution method for estimation of flash point, *Energy Fuels* 22 (2008) 1628–1635.
- [13] F. Gharagheizi, A new neural network quantitative structure-property relationship for prediction of θ (Lower Critical Solution Temperature) of Polymer Solutions, *e-Polymers* (2007), Article number 114.
- [14] F. Gharagheizi, A. Fazeli, Prediction of Watson characterization factor of hydrocarbon compounds from their molecular properties, *QSAR Comb. Sci.* 27 (2008) 758–767.
- [15] F. Gharagheizi, R.F. Alamdari, A molecular-based model for prediction of solubility of c60 fullerene in various solvents, *Fuller. Nanotub. Car. N.* 16 (2008) 40–57.
- [16] F. Gharagheizi, M. Mehrpooya, Prediction of some important physical properties of sulfur compounds using QSPR models, *Mol. Divers.* 12 (2008) 143–155.
- [17] F. Gharagheizi, B. Tirandazi, R. Barzin, Estimation of aniline point temperature of pure hydrocarbons: a quantitative structure-property relationship approach, *Ind. Eng. Chem. Res.* 48 (2009) 1678–1682.
- [18] M. Sattari, F. Gharagheizi, Prediction of molecular diffusivity of pure components into air: A QSPR approach, *Chemosphere* 72 (2008) 1298–1302.
- [19] F. Gharagheizi, Prediction of standard enthalpy of formation of pure compounds using molecular structure, *Aus. J. Chem.* 62 (2009) 374–381.
- [20] T. Albahri, R.S. George, Artificial neural network investigation of structural group contribution method for predicting pure components auto ignition temperature, *Ind. Eng. Chem. Res.* 42 (2003) 5708–5714.
- [21] T. Albahri, Structural group contribution method for predicting the octane number of pure hydrocarbon liquids, *Ind. Eng. Chem. Res.* 42 (2003) 657–662.
- [22] F. Gharagheizi, QSPR analysis for intrinsic viscosity of polymer solutions by means of GA-MLR and RBFNN, *Comput. Mater. Sci.* 40 (2007) 159–167.