Thermochimica Acta 512 (2011) 150–156

Contents lists available at ScienceDirect

Thermochimica Acta

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

Neural network/particle swarm method to predict flammability limits in air of organic compounds

Juan A. Lazzús*

Departamento de Física, Universidad de La Serena, Casilla 554, La Serena, Chile

ARTICLE INFO

Article history:
Received 29 June 2010
Received in revised form 28 July 2010
Accepted 28 September 2010
Available online 8 October 2010

Keywords:
Flammability
Lower flammability limit percentage
Upper flammability limit percentage
Artificial neural networks
Particle swarm optimization
Group contribution method

ABSTRACT

A new neural network based on particle swarm optimization was developed to estimate lower and upper flammability limit percent in air of organic compounds using their molecular structures. The capability of this hybrid method was tested using 328 compounds in a correlated set and 90 compounds in a predicted set. The results show that the proposed method can estimate the flammability limits of organic compounds with low deviations and can be used with acceptable accuracy in thermal engineering.

1. Introduction

The thermal properties used to determine the potential for fire and explosion hazards of industrial substances are the flammability limits. To safe handling, transportation, and storage of flammable compounds, information about flammability of these substances is needed [1]. The most important properties used to his purpose are the lower flammability limit percent (*LFLP*) and the upper flammability limit percent (*UFLP*). Flammability limits provide range of fuel concentration (normally in percentage volume), within which a gaseous mixture can ignite and burn [2]. The lower flammable limit is the smallest percent of the gas in air which can ignite when exposed to the ignition temperature. The upper flammable limit is the point above which the mixture is too rich in fuel to ignite. The range between these two limits is the flammable or explosive range [3].

Knowledge of *LFLP* and *UFLP* of several substances is essential in many thermal and chemical engineering operations. A precise determination of flammability limits in air requires the use of standard apparatus and conditions. When the flammability limits cannot be determined experimentally, other methods of estimation are necessary. Jones [4] proposed equations frequently used:

$$LFLP = 0.55C_{\mathrm{est}} \quad (1)$$

$$UFLP = 3.50C_{\mathrm{est}} \quad (2)$$

where C_{est} is the stoichiometric concentration of the flammable product for complete combustion in air. Another empirical relation used for the prediction of *LFLP* in air was proposed by Spakowski [5]:

$$LFLP = -\frac{4354}{\Delta H_{\mathrm{comb}}} \quad (3)$$

where ΔH_{comb} is the heat of combustion in kJ mol^{-1}. Hristova et al [6] developed a correlation of flammability limits with normal boiling point and molecular mass. Recently, quantitative structure–property relationship (QSPR) methods were presented by several authors for the prediction of *LFLP* [7–10] and *UFLP* [7,11,12] of pure compounds.

There are many methods for the prediction of thermal properties in the literature [1]. Artificial neural networks (ANN) are accepted as the most powerful non-linear technique in several applications [13]. The neural network modeling has been applied to most physicochemical properties, for which suitable experimental data can be found in the literature. Alternatives of predictive methods have recently appeared. The group contribution method (GCM) provides a convenient tool to predict physicochemical properties of substances from molecular structural information [14]. The mentioned GCM use linear and non-linear regression techniques to represent the relations among the variables of a given system [15]. The relationship between the physical and thermodynamic properties is highly non-linear, and consequently an ANN

---

* Corresponding author. Tel.: +56 51 204128; fax: +56 51 206658.
E-mail address: jlazzus@dfuls.cl.

0040-6031/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.tca.2010.09.018

can be a suitable alternative to model the underlying thermophysic properties [13–15].

In this work, the flammability limits (*LFLP* and *UFLP*) in air of organic compounds have been estimated using a simple GCM implemented in an ANN replacing standard back-propagation with particle swarm optimization (PSO) [16], that is one of the most recently developed evolutionary algorithms.

2. Neural network and particle swarm optimization

A feed-forward neural network programmed with the software MatLab [17], was used to represent non-linear relationships among variables [13–15]. This ANN program considers the reading of the necessary data organized in an Excel file. To distinguish between the different physical and chemical properties of the substances used, and so the network can discriminate and learn in optimum form, properties derived from the molecular structure were considered. The input layer contains one neuron (node) for each variable. The output layer has one node generating the scaled estimated value of the flammability limits. The ANN was trained with particle swarm optimization [14].

PSO is a population-based optimization tool, where the system is initialized with a population of random particles and the algorithm searches for optima by updating generations [15]. In each iteration, the velocity for each particle is calculated according to the following formula [16]:

$$v_i^p(t+1) = \omega v_i^p(t) + c_1 r_1(\psi_i^p(t) - x_i^p(t)) + c_2 r_2(\psi_g(t) - x_i^p(t)) \qquad (4)$$

where t is the current step number, ω is the inertia weight, c_1 and c_2 are the acceleration constants, and r_1, r_2 are element from two randon sequences in the range (0,1). $x_i^p(t)$ is the current position of the particle, ψ_i^p is the best one of the solutions that this particle has reached, and ψ_g is the best solutions that all the particles have reached. In general, the value of each component in v can be clamped to the range $(-v_{\max}, v_{\max})$ control excessive roaming of particles outside the search space [14,15]. After calculating the velocity, the new position of every particle is:

$$x^p(t+1) = x^p(t) + v^p(t+1) \qquad (5)$$

when x and v denote a particle position and its corresponding velocity in a search space, respectively.

The total steps to calculate the output parameter, using the input parameters, were as follows [13]:

The data were normalized using the following equation:

$$p_i = (P_i - P_i^{\min})\frac{2}{P_i^{\max} - P_i^{\min}} - 1 \qquad (6)$$

where P_i is the input variables i, $P_i^{\min}$ and $P_i^{\max}$ are the smallest and largest value of the data. Next, the net inputs (N) are calculated for the hidden neurons coming from the inputs neurons. For a hidden neuron:

$$N_j^h = \sum_i^n w_{ij}^h p_i + b_j^h \qquad (7)$$

where p_i is the vector of the inputs of the training, j is the hidden neuron, w_{ij} is the weight of the connection among the input neurons with the hidden layer, and the term b_j corresponds to the bias of the neuron j of the hidden layer, reached in its activation. Starting from these inputs, the outputs (y) of the hidden neurons are calculated, using a transfer function f^h associated with the neurons of this layer.

$$y_j^h = f_j^h\left(\sum_i^n w_{ij}^h p_i + b_j^h\right) \qquad (8)$$

To minimize the error, the transfer function f it should be differentiable. In the ANN, the hyperbolic tangent function (tansig) was used as:

$$f(N_{jk}) = \frac{e^{N_{jk}} - e^{-N_{jk}}}{e^{N_{jk}} + e^{-N_{jk}}} \qquad (9)$$

All the neurons of the ANN have an associated activation value for a give input pattern, the algorithm continues finding the error that is presented for each neuron, except those of the input layer. After finding the output values, the weights of all layers of the network are actualized by PSO, using Eqs. (4) and (5) [14].

The PSO algorithm is very different then any of the traditional methods of training [15]. Each neuron contains a position and velocity. The position corresponds to the weight of a neuron. The velocity is used to update the weight. The velocity is used to control how much the position is updated. On each pass through a data set, PSO compares each weight's fitness. The network with the highest fitness is considered the global best. The other weights are updated based on the global best network rather than on their personal error or fitness [14–16]. In this study, network fitness was determined to be the mean square of errors for the entire training set:

$$\text{Fitness} = \sum (\text{Recorded\_Value} - \text{Network\_Predicted\_Value})^2 \qquad (10)$$

This process repeats for the total number of patterns to training, for a successful process the objective of the algorithm is to modernize all the weights minimizing the total mean squared error:

$$E^2 = \frac{1}{2}\sum (\text{Fitness})^2 \qquad (11)$$

Fig. 1 presents a block diagram of the program developed and written in MatLab M-file [17].

3. Database used and training

A heterogeneous set of compounds: aromatic and aliphatic hydrocarbons, halogens, polychlorinated biphenyls, mercaptans, sulfides, anilines, pyridines, alcohols, carboxylic acids, aldehydes, amines, ketones, and esters, were included. A total of 328 substances were used to train the network and then values of *UFLP* and *LFLP* of 90 substances not used in the training process, were predicted (the total set of substances is available as supplementary data). To distinguish between the different substances considered in this study, so the net can discriminate and learn in optimum form, the properties used cover wide ranges: 0.2–15.7 (vol% in air) for *LFLP*, and 1.8–60.0 (vol% in air) for *UFLP*. Molecular mass M (size), dipole moment μ (polarity), and the structure of the molecules represented by the number of well defined groups forming the molecule, were provided as input variables. Molecular mass and dipole moment were chosen to characterize the different molecules [13–15]. In addition, the substances included in the study have very different physical and chemical characteristics. Low molecular weight substances such as ethylene ($M = 28$) to high molecular weight substances such as hexatriacontane ($M = 506$); or non-polar substances ($\mu = 0$) such as benzene, to highly polar substances such as dinitrobenzene ($\mu = 6.3$). Thus, the problem is not straightforward and probably is one of the reasons why the flammability limits have not been treated as proposed in this work. Note that, all the properties of interest (M, μ, *LFLP*, *UFLP*) were taken from the DIPPR database [18]. Fig. 2 shows a general picture of the range of flammability limits in air and substances considered in this study.

This work used a leave-20%-out cross-validation method to estimate the predictive capabilities of the model. Training and prediction sets were selected randomly, with the consideration that in the group contribution methods, the molecules are decomposed into fragments and all fragments that are present with adequate frequency in the training database. The structure of the molecules

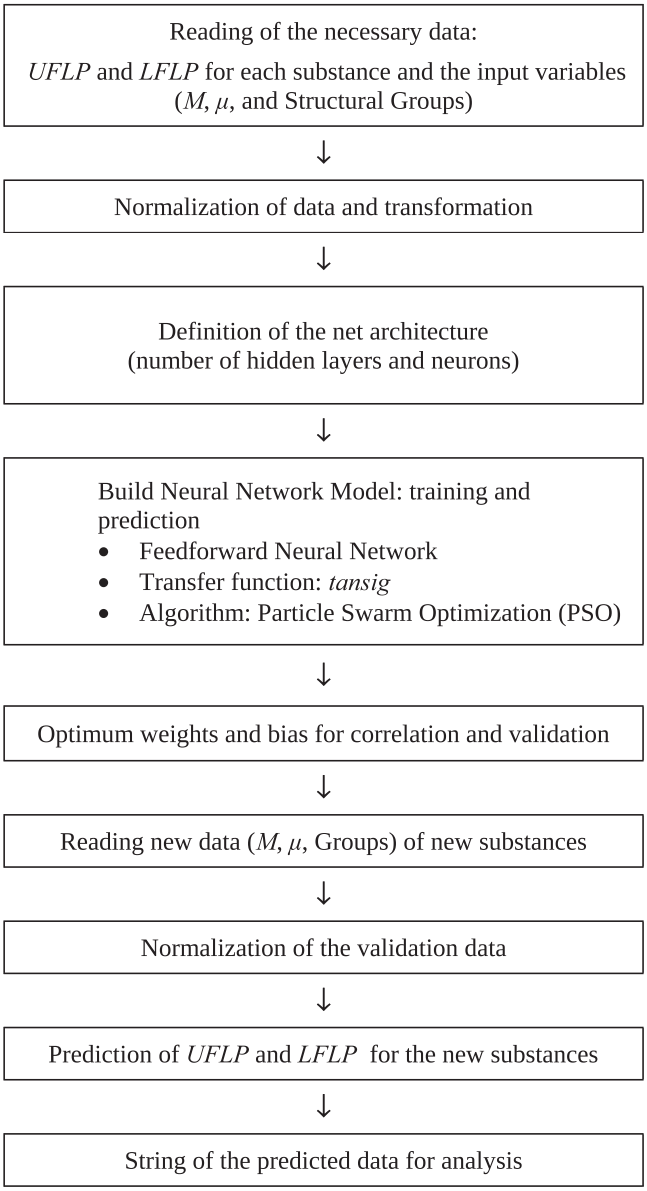

Fig. 1. Flow diagram for the ANN + PSO program developed for this work.

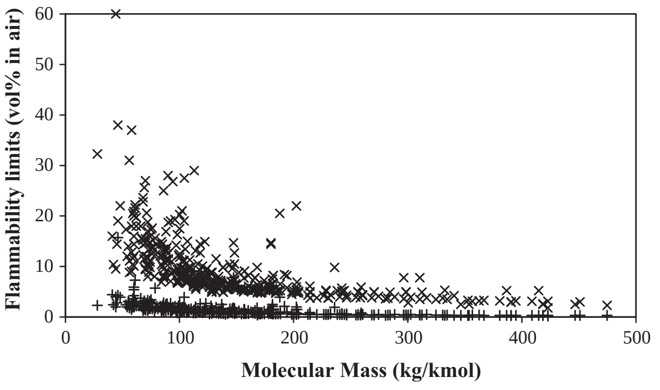

Fig. 2. Flammability limits as a function of molecular mass for all the substances used in this study, and taken from Ref. [14]: *LFLP* (+) and *UFLP* (×).

Table 1
Structural groups used in the proposed GCM + ANN + PSO model.

Parameter			No. occurrence		
No.	Group	Max. value	Training set	Perdiction set	Total set
1	*M*	506.98	328	90	418
2	μ	6.30	261	74	335
3	$-CH_3$	8	226	60	286
4	$-CH_2-$	34	173	43	216
5	>CH–	6	37	1	38
6	>C<	2	14	3	17
7	$=CH_2$	2	28	4	32
8	=CH–	2	26	5	31
9	=C<	1	12	1	13
10	=C=	1	3	0	3
11	(CH	1	2	0	2
12	(C–	2	2	0	2
13	–OH	5	29	8	37
14	–O–	2	13	0	13
15	>C=O	2	13	5	18
16	–CHO	2	28	3	31
17	–COOH	2	28	7	35
18	–COO–	3	26	6	32
19	HCOO–	1	3	3	6
20	=O	1	2	0	2
21	$-NH_2$	2	14	5	19
22	–NH–	2	11	3	14
23	>N–	1	5	1	6
24	–CN	2	10	3	13
25	$-NO_2$	2	9	5	14
26	–F	2	2	0	2
27	–Cl	3	11	5	16
28	–Br	2	7	1	8
29	–SH	1	12	1	13
30	–S–	2	5	3	8
31	$-CH_2-$(ring)	10	27	5	32
32	>CH–(ring)	6	12	1	13
33	=CH–(ring)	20	120	48	168
34	>C<(ring)	2	6	0	6
35	=C<(ring)	6	117	43	160
36	–O–(ring)	1	8	4	12
37	–OH(ring)	6	18	11	29
38	>C=O(ring)	2	8	1	9
39	–NH–(ring)	1	3	3	6
40	>N–(ring)	1	2	1	3
41	=N–(ring)	1	12	1	13
42	–S–(ring)	1	2	2	4

used, was represented by the number of well-defined groups forming the molecules. The value associated to the structural group was defined as following: 0, when the group does not appear in the substance, and *n*, when the group appears *n*-times in the substance [13,15]. For instance, for ethyl vanillin ($C_9H_{10}O_3$, CASN: 121-32-4), the property data are $M = 166.2\ \mathrm{kg\ kmol^{-1}}$, $\mu = 4.2$ debyes, and the structure of the molecule: $[-CH_3] = 1$, $[-CH_2-] = 1$, $[-O-] = 1$, $[-CHO] = 1$, $[=CH-(ring)] = 3$, $[=C<(ring)] = 3$, and $[-OH(ring)] = 1$. Table 1 shows the 42 groups used as entrance variables.

Several network architectures were tested to select the most accurate scheme. The most basic architecture normally used for this type of application involves a neural network consisting of three layers [13]. The number of hidden neurons needs to be sufficient to ensure that the information contained in the data utilized for training the network is adequately represented. There is no specific approach to determine the number of neurons of the hidden layer, many alternative combinations are possible. The optimum number of neurons was determined by adding neurons in systematic form and evaluating the average absolute deviations of the sets during the learning process [14,16]. Fig. 3 shows the average absolute relative deviation found in correlating the flammability limits in air of all compounds as a function of the number of neurons in the hidden layer. The network that gave the lowest deviation during training was one with 42 parameters in the input layer, 6 neurons

Table 2
Substances used in the prediction set, and deviations obtained using the GCM + ANN + PSO method.

Prediction set no.	*LFLP* (vol% in air)				*UFLP* (vol% in air)				
	Formula	CASN	Substance	exp	calc	Δ%	exp	calc	Δ%
1	CH_3NO_2	75-52-5	Nitromethane	7.3	7.1	2.2	22.2	22.3	0.6
2	CH_4S	74-93-1	Methyl mercaptan	3.9	3.9	0.4	22.0	22.7	3.0
3	$C_2H_4O_2$	107-31-3	Methyl formate	5.9	5.4	9.2	20.0	19.8	1.0
4	C_3H_3N	107-13-1	Acrylonitrile	2.4	2.5	3.9	17.3	18.5	6.5
5	C_3H_4O	107-02-8	Acrolein	2.8	2.6	8.2	31.0	31.4	1.1
6	C_3H_5N	107-12-0	Propionitrile	3.1	2.6	16.3	14.0	12.7	9.0
7	C_3H_5NO	79-06-1	Acrylamide	2.7	2.7	0.4	20.6	20.4	1.0
8	C_3H_6	75-19-4	Cyclopropane	2.4	2.3	3.3	10.4	11.3	9.0
9	C_3H_6O	75-56-9	1,2-Propylene oxide	1.9	1.9	2.2	37.0	36.8	0.6
10	C_3H_6S	287-27-4	Trimethylene sulfide	2.0	1.9	2.6	15.9	14.0	12.0
11	C_3H_8S	624-89-5	Methyl ethyl sulfide	1.8	1.7	5.8	13.9	14.0	0.6
12	$C_4H_4Cl_2$	106-46-7	*p*-Dichlorobenzene	1.8	1.7	3.7	7.8	7.7	1.8
13	C_4H_4S	110-02-1	Thiophene	1.6	1.6	1.1	14.9	14.2	4.7
14	C_4H_6O	78-85-3	Methacrolein	2.1	1.9	8.4	14.6	14.0	3.8
15	C_4H_7NO	75-86-5	Acetone cyanohydrin	2.2	1.8	19.7	12.0	11.6	3.2
16	C_4H_8O	78-93-3	Methyl ethyl ketone	1.8	1.5	14.6	10.0	10.8	7.9
17	C_4H_8OS	3268-49-3	3-Methylmercapto propanal	1.6	1.7	9.2	27.5	26.0	5.3
18	C_4H_9N	123-75-1	Pyrrolidine	2.4	2.1	12.2	12.0	14.1	17.6
19	C_4H_9NO	110-91-8	Morpholine	1.8	1.7	4.5	10.8	10.4	3.6
20	$C_4H_{10}S$	3877-15-4	Methyl n-propyl sulfide	1.4	1.3	6.6	11.5	11.5	0.2
21	$C_4H_{11}N$	75-64-9	*tert*-Butylamine	1.7	1.4	17.3	8.9	9.3	4.4
22	$C_4H_{11}NO_2$	111-42-2	Diethanolamine	1.8	1.9	4.1	13.4	12.2	9.3
23	C_5H_5N	110-86-1	Pyridine	1.8	1.6	12.9	12.4	9.8	21.0
24	C_5H_7N	96-54-8	*N*-Methylpyrrole	1.6	1.7	5.9	14.0	15.4	9.7
25	$C_5H_{10}O_2$	542-55-2	Isobutyl formate	1.7	1.6	8.3	8.0	8.7	8.7
26	$C_5H_{10}O_2$	105-37-3	Ethyl propionate	1.9	1.6	16.0	11.0	9.6	12.4
27	$C_5H_{10}O_2$	592-84-7	*n*-Butyl formate	1.7	1.5	9.7	8.0	10.2	27.8
28	$C_5H_{11}N$	110-89-4	Piperidine	1.4	1.7	20.7	10.0	11.2	12.2
29	$C_6H_3Cl_3$	120-82-1	1,2,4-Trichlorobenzene	2.5	2.3	7.7	6.6	7.5	13.2
30	$C_6H_3Cl_3$	108-70-3	1,3,5-Trichlorobenzene	2.1	2.3	9.5	7.5	8.5	13.1
31	$C_6H_3ClN_2O_4$	97-00-7	1-Chloro-2,4-dinitrobenzene	2.0	2.4	20.2	22.0	19.8	10.1
32	C_6H_5Br	108-86-1	Bromobenzene	1.5	1.3	11.6	9.1	8.3	9.0
33	C_6H_5Cl	108-90-7	Monochlorobenzene	1.3	1.1	13.0	7.1	8.8	24.2
34	C_6H_6	71-43-2	Benzene	1.2	1.2	0.1	8.0	7.9	1.5
35	$C_6H_6N_2O_2$	88-74-4	*o*-Nitroaniline	1.5	1.5	2.9	9.8	10.4	5.9
36	$C_6H_6N_2O_2$	99-09-2	*m*-Nitroaniline	1.7	1.6	6.8	9.8	10.6	7.7
37	$C_6H_6N_2O_2$	100-01-6	*p*-Nitroaniline	1.5	1.6	9.5	9.8	10.8	10.2
38	C_6H_6O	108-95-2	Phenol	1.5	1.3	10.2	9.1	8.5	6.3
39	$C_6H_6O_2$	120-80-9	1,2-Dihydroxybenzene	1.6	1.6	2.6	9.8	9.3	5.6
40	$C_6H_{10}O_4$	124-04-9	Adipic acid	1.6	1.6	2.8	9.6	10.0	3.7
41	$C_6H_{12}O_2$	142-62-1	Caproic acid	1.3	1.3	0.7	8.2	8.5	3.7
42	C_7H_8O	108-39-4	*m*-Cresol	1.1	1.2	9.7	7.6	7.8	2.7
43	C_7H_8O	106-44-5	*p*-Cresol	1.1	1.2	9.8	7.6	7.8	2.7
44	C_7H_8O	95-48-7	*o*-Cresol	1.4	1.2	13.7	7.6	7.8	2.8
45	$C_8H_4O_3$	85-44-9	Phthalic anhydride	1.7	1.6	7.3	10.5	9.8	7.1
46	C_8H_8O	98-86-2	Acetophenone	1.1	1.0	12.2	6.7	7.6	12.9
47	C_8H_9NO	103-84-4	Acetanilide	1.1	0.9	15.0	6.9	6.9	0.0
48	C_8H_{10}	106-42-3	*p*-Xylene	1.1	1.0	9.0	7.0	6.7	4.5
49	C_8H_{10}	95-47-6	*o*-Xylene	0.9	1.0	10.7	6.7	6.7	0.7
50	C_8H_{10}	108-38-3	*m*-Xylene	1.1	1.0	9.2	7.0	6.7	4.7
51	$C_8H_{10}O$	105-67-9	2,4-Dimethylphenol	1.1	1.1	0.7	6.4	7.2	12.3
52	$C_8H_{10}O$	108-68-9	3,5-Dimethylphenol	1.1	1.1	0.9	6.4	7.2	12.0
53	$C_8H_{10}O$	95-65-8	3,4-Dimethylphenol	1.1	1.1	0.9	6.4	7.2	12.0
54	$C_8H_{10}O$	576-26-1	2,6-Dimethylphenol	1.4	1.1	22.0	6.4	7.2	12.3
55	$C_8H_{10}O$	95-87-4	2,5-Dimethylphenol	1.1	1.1	0.7	6.4	7.2	12.3
56	$C_8H_{10}O$	526-75-0	2,3-Dimethylphenol	1.1	1.1	0.6	6.4	7.2	12.4
57	C_9H_{12}	103-65-1	*n*-Propylbenzene	0.9	0.9	0.9	6.0	6.0	0.4
58	$C_9H_8O_2$	140-10-3	Cinnamic acid	1.0	1.0	2.7	7.4	8.2	10.5
59	$C_{10}H_{11}NO_2$	102-01-2	Acetoacetanilide	0.9	0.7	24.1	6.1	7.2	18.3
60	$C_{10}H_{12}O_2$	2315-68-6	*n*-Propyl benzoate	0.9	0.9	0.9	5.6	6.2	10.9
61	$C_{10}H_{14}$	104-51-8	*n*-Butylbenzene	0.8	0.8	2.1	5.8	5.6	3.7
62	$C_{10}H_{15}N$	91-66-7	*N*,*N*-Dimethylaniline	0.8	0.8	3.6	6.4	5.8	9.8
63	$C_{10}H_{18}O_4$	111-20-6	Sebacic acid	0.8	0.9	13.6	6.0	6.3	5.2
64	$C_{10}H_{20}O_2$	334-48-5	*n*-Decanoic acid	0.8	0.8	3.8	5.5	5.7	4.0
65	$C_{10}H_{24}N_2$	4062-60-6	*N*,*N*'-Di-*tert*-Butylethylenediamine	0.7	0.7	5.9	4.9	5.2	6.7
66	$C_{11}H_{14}O_2$	136-60-7	*n*-Butyl benzoate	0.8	0.8	3.2	5.4	5.7	5.4
67	$C_{11}H_{16}$	538-68-1	*n*-Pentylbenzene	0.8	0.8	5.1	5.5	5.2	4.9
68	$C_{12}H_8O$	132-64-9	Dibenzofuran	0.8	0.8	5.4	6.4	6.7	4.3
69	$C_{12}H_{26}O$	112-53-8	1-Dodecanol	0.6	0.7	13.8	5.1	4.8	6.1
70	$C_{14}H_{12}$	103-30-0	*trans*-Stilbene	0.7	0.5	26.3	5.3	5.6	6.6
71	$C_{14}H_{14}$	103-29-7	Bibenzyl	0.6	0.7	12.5	5.2	4.7	9.6
72	$C_{14}H_{30}O$	112-72-1	1-Tetradecanol	0.5	0.6	21.7	5.0	4.3	13.3
73	$C_{16}H_{34}O$	36653-82-4	1-Hexadecanol	0.5	0.6	11.1	5.0	4.0	19.6
74	$C_{17}H_{28}$	6742-54-7	*n*-Undecylbenzene	0.6	0.6	4.8	4.4	4.1	7.2

Table 2 (*Continued*)

Prediction set no.	*LFLP* (vol% in air)				*UFLP* (vol% in air)				
	Formula	CASN	Substance	exp	calc	Δ%	exp	calc	Δ%
75	$C_{17}H_{36}O$	1454-85-9	1-Heptadecanol	0.4	0.4	8.2	5.0	3.9	22.2
76	$C_{18}H_{30}$	123-01-3	*n*-Dodecylbenzene	0.6	0.6	7.9	4.3	4.0	7.5
77	$C_{18}H_{32}O_2$	57-11-4	Stearic acid	0.4	0.4	1.3	4.9	4.8	2.7
78	$C_{18}H_{38}O$	112-92-5	1-Octadecanol	0.4	0.5	28.2	5.0	3.8	24.6
79	$C_{19}H_{38}O_2$	646-30-0	Nonadecanoic acid	0.4	0.4	9.7	4.9	4.9	0.7
80	$C_{20}H_{30}O_4$	84-75-3	Di-*n*-Hexyl phthalate	0.4	0.5	12.7	3.5	3.5	1.2
81	$C_{20}H_{34}$	1459-10-5	*n*-Tetradecylbenzene	0.5	0.5	3.9	4.1	3.8	7.6
82	$C_{20}H_{42}O$	629-96-9	*n*-Eicosanol	0.4	0.5	19.6	5.0	4.6	8.5
83	$C_{21}H_{36}$	2131-18-2	*n*-Pentadecylbenzene	0.5	0.5	0.8	4.0	3.7	7.4
84	$C_{22}H_{34}O_4$	3648-21-3	Di-*n*-Heptyl phthalate	0.4	0.4	1.7	3.2	3.2	0.2
85	$C_{25}H_{52}$	629-99-2	*n*-Pentacosane	0.3	0.3	10.1	3.3	3.3	0.0
86	$C_{26}H_{54}$	630-01-3	*n*-Hexacosane	0.3	0.3	5.3	3.3	3.2	2.4
87	$C_{27}H_{56}$	593-49-7	*n*-Heptacosane	0.3	0.3	6.9	3.2	3.1	1.8
88	$C_{28}H_{46}O_4$	84-77-5	Di-*n*-Decyl phthalate	0.3	0.3	12.7	2.5	2.8	11.2
89	$C_{28}H_{58}$	630-02-4	*n*-Octacosane	0.3	0.3	3.0	3.2	3.1	4.1
90	$C_{29}H_{60}$	630-03-5	*n*-Nonacosane	0.3	0.4	25.3	3.1	3.0	3.4

Table 3
Overall minimum, maximum, and average deviations for the calculated flammability limits of all compounds using the proposed method.

Statistics	*LFLP* (vol% in air)			*UFLP* (vol% in air)		
	Training set	Prediction set	Total set	Training set	Prediction set	Total set
No. of substances	328	90	418	328	90	418
$\Delta\%_{\min}$	0.0	0.0	0.0	0.0	0.0	0.0
$\Delta\%_{\max}$	26.1	28.2	28.2	26.8	27.8	27.8
Δ%<10	204	60	264	238	64	302
Δ%>20	17	8	25	11	5	16
AARD	8.6	8.5	8.6	7.0	7.5	7.1
R^2	0.9876	0.9819	0.9865	0.9780	0.9771	0.9818

in the hidden layer, and two neurons in the output layer. For this architecture the average deviation during training is 8.6% for *LFLP* and 7.0% for *UFLP*, and during prediction is 8.5% for *LFLP* and 7.5% for *UFLP*.

The accuracy of the chosen final network was checked using the average absolute deviation (AARD) between the calculated value of *LFLP* and *UFLP* after training and the data from the literature. The deviations were calculated as:

$$\Delta\% = \left|\frac{\text{calc} - \text{exp}}{\text{exp}}\right| 100 \tag{12}$$

$$AARD_{LFLP} = \frac{100}{N_D}\sum_{i=1}^{N}\left|\frac{LFLP_{\text{calc}} - LFLP_{\text{exp}}}{LFLP_{\text{exp}}}\right|_i \tag{13}$$

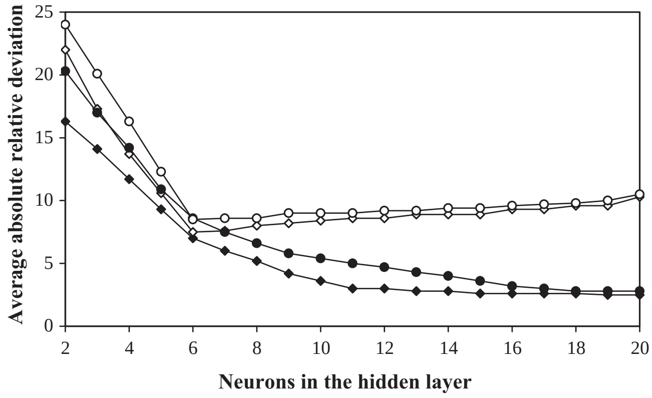

Fig. 3. Average absolute relative deviations found in correlating the flammability limits of all substances as function of the number of neurons in the hidden layer during the training step for *LFLP* (●) and *UFLP* (♦), and during the prediction step for *LFLP* (○) and *UFLP* (◊).

$$AARD_{UFLP} = \frac{100}{N_D}\sum_{i=1}^{N}\left|\frac{UFLP_{\text{calc}} - UFLP_{\text{exp}}}{UFLP_{\text{exp}}}\right|_i \tag{14}$$

4. Results and discussion

Table 2 shows the detail of the 90 substances used in the prediction set and the deviations obtained. Table 3 shows the overall minimum, maximum, and average deviations for all the substances using the proposed ANN 42-6-2. The results show that the ANN can be accurately trained and that the chosen architecture can estimate *LFLP* and *UFLP* of organic compounds with enough accuracy. This new method, gives lower deviations than others method yet available in the literature: AARD less than 8.6% for *LFLP* and 7.0% for *UFLP* in the training step, and AARD less than 8.5% for *LFLP* and 7.5% for *UFLP* in the prediction step.

Once the best architecture was determining, the optimum weights required to carry out the estimate of flammability limits of organic compounds, were obtained. Table 4 shows the optimum weights and biases for the ANN 42-6-2.

Fig. 4 shows a comparison between experimental (solid line) and calculated values of *LFLP*. In the training step (×) the correlation coefficient (R^2) was 0.9876, and during the prediction step (○) R^2 was 0.9819. Fig. 5 shows a similar comparison for the calculated values of *UFLP*. In this case, R^2 was 0.9780 for the training set, and R^2 was 0.9771 for the prediction set.

Fig. 6 shows a comparison between several methods proposed in the literature for prediction of *LFLP* for diverse organic compounds [7–10], and the GCM + ANN + PSO model proposed in this work. As see in this figure, Pan et al. [7] shows *LFLP* in a range of 0.1–2.3 vol% in air, with $\Delta\%_{\max}$ higher than 35%. Pan et al. [8] shows *LFLP* in a range of 0.2–3.6 vol% in air, with AARD of 5.6% and $\Delta\%_{\max}$ greater than 60%. Gharagheizi [9] shows *LFLP* in a range of 0.1–4.3 vol% in air, with AARD of 7.6% and $\Delta\%_{\max}$ greater than 50%. Gharagheizi method [10] shows *LFLP* in a range of 0.1–4.3 vol% in air, with

Table 4
Optimum weights and biases of the model 42-6-2.

w_{ji}	1	2	3	4	5	6
1	0.1557	−0.0440	−5.6480	−1.3415	7.5607	−1.0829
2	−0.1541	−0.7115	0.0072	0.2745	0.0435	−0.0536
3	0.3991	−0.2208	−0.7054	−0.1312	0.9305	−0.3508
4	−2.6097	0.5296	−3.0392	0.0738	4.1433	−0.9456
5	2.0721	−0.3348	−2.4600	−0.1413	2.6844	−0.0275
6	−0.5279	−0.7798	−1.2530	−0.0054	1.3230	−0.0569
7	−0.8413	−0.4538	−0.0734	1.0779	−0.1129	−0.1745
8	3.1582	0.8776	−0.2432	−0.7556	0.4463	−0.0995
9	−0.9341	0.5735	0.0348	0.5624	−0.1691	0.0650
10	−3.6402	−0.0605	−0.2793	−0.1662	−0.1365	0.1953
11	−0.8294	−0.6220	0.2516	0.0011	−0.9245	0.1321
12	−1.0662	0.0191	0.5303	0.3406	−0.6100	0.0068
13	1.8597	0.1713	1.1153	−0.6923	−1.9963	0.4548
14	1.8981	0.0494	−1.3560	−1.1539	0.2328	0.6335
15	2.7299	1.8691	0.1056	−0.7002	−0.1582	−0.1024
16	−1.3606	−1.0969	0.8563	−0.3081	−1.4723	0.1072
17	−0.0961	0.5613	0.7458	−0.2403	−0.9773	0.2846
18	1.0590	−0.6046	1.5035	−0.2929	−2.2809	0.4568
19	0.3849	−0.0764	−0.3307	−0.3345	0.8193	0.2115
20	0.5278	−0.7928	2.0456	1.0196	−0.4295	−1.5149
21	0.3334	−0.1205	0.3177	−0.5472	−0.4567	0.2129
22	−0.4519	0.3912	0.0157	−0.1623	−0.2019	0.2198
23	−1.1981	0.3289	0.0183	−0.2924	−0.1208	0.1508
24	0.7382	0.3944	0.0735	−0.3245	−0.1256	0.2575
25	0.7707	−0.0753	0.8177	−0.8328	−1.0172	0.6240
26	−0.4737	−0.1443	0.6436	−0.0209	−0.8470	0.2505
27	4.9148	0.0903	1.6839	1.6860	−1.7196	−0.7024
28	0.8132	−1.9150	2.0325	1.0031	1.1603	0.3627
29	−1.6827	0.5317	0.1378	−0.7642	−0.3016	0.4213
30	−1.2851	−4.2418	0.6075	−0.7107	−1.0689	0.0113
31	−1.3052	1.4870	0.1834	−0.1898	−0.6631	−0.3225
32	0.9210	−0.7426	0.0474	0.6940	−1.9099	−0.3290
33	0.0084	0.2673	−1.9888	−0.0298	2.1500	−0.5692
34	0.9309	0.2156	0.1092	−0.2546	−0.3471	0.0320
35	−0.5351	0.3670	0.3479	0.0905	0.4843	−0.0072
36	1.5814	−1.9008	0.2511	0.7810	−0.2659	−0.1551
37	−0.0214	0.0160	1.3234	−0.4979	−0.2791	0.4417
38	−2.0619	1.3683	0.6107	0.1132	−0.8968	−0.1773
39	−2.3615	0.6794	0.0326	0.0616	−0.1173	0.1385
40	0.0887	−0.0334	0.5587	0.0755	−0.6613	−0.3212
41	−0.4365	0.9536	0.4781	0.4693	−0.7593	0.0626
42	0.0109	−0.7098	0.3728	−0.0286	−0.5141	−0.3929
b_j	−1.3099	0.9280	−0.3684	−0.3846	1.6389	−1.2494
1	0.1480	−0.0943	4.5340	0.4811	3.3038	0.5831
2	0.2901	−0.4075	−0.4207	0.4060	−4.2936	0.6229
b_k	−0.2428	3.2058				

AARD of 4.6% and $\Delta\%_{max}$ higher than 50%. And the GCM + ANN + PSO method shows AARD of 8.6% and $\Delta\%_{max}$ lower than 30%, for a range of 0.2–15.7 vol% in air.

Fig. 7 shows a comparison between several QSPR methods proposed in the literature to predict of *UFLP* of organic compounds [7,11,12], and the GCM + ANN + PSO model proposed in this work. As see in this figure, Pan et al. [7] shows *UFLP* in a range of 3.0–16.5 vol% in air, with $\Delta\%_{max}$ greater than 40%. QSPR method [11] shows *UFLP* with AARD of 19.2% and $\Delta\%_{max}$ higher than 100% for a range of 2.7–40.3 vol% in air. Gharagheizi [12] shows *UFLP* with AARD of 9.6% and $\Delta\%_{max}$ higher than 30%, on a range of 2.3–23.5 vol% in air. And the GCM + ANN + PSO method shows AARD of 7.1% and $\Delta\%_{max}$ lower than 30%, for a range of 1.8–60.0 vol% in air.

Another comparison was made with a neural network with standard backpropagation (BPNN), and similars architecture (42-6-2) and database. This BPNN show results of AARD of 15%, and $\Delta\%_{max}$

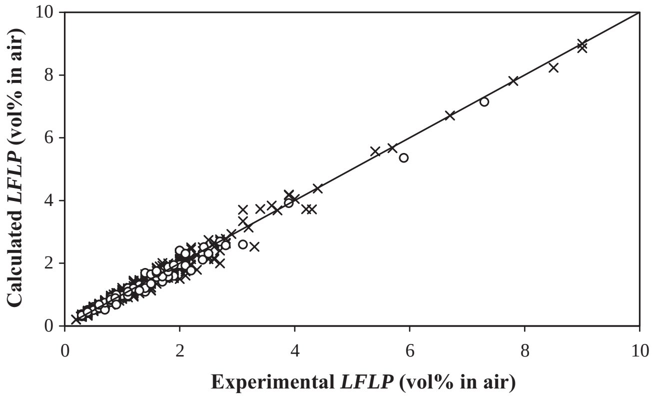

Fig. 4. Comparison between experimental and calculated values of lower flammability limit percent in air: (×) during the training and (○) during the prediction.

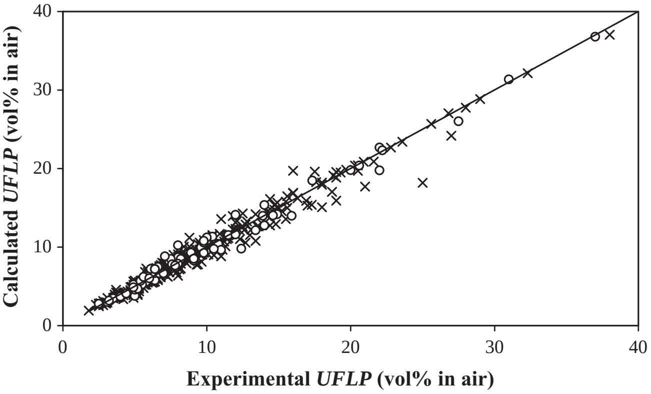

Fig. 5. Comparison between experimental and calculated values of upper flammability limit percent in air: (×) during the training and (○) during the prediction.

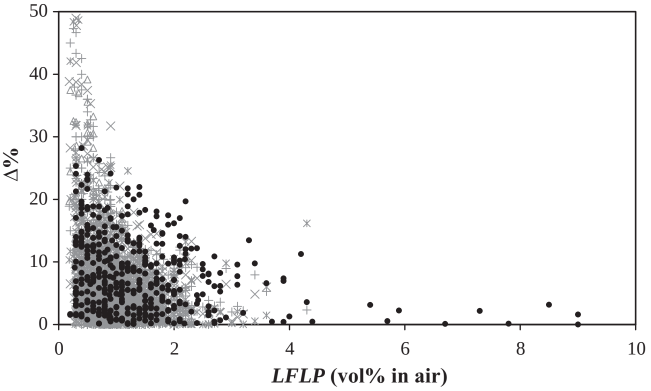

Fig. 6. Average absolute relative deviations found in the prediction of lower flammability limit percent in air using: (Δ) Pan et al. [7], (×) Pan et al. [8], (+) Gharagheizi [9], (*) Gharagheizi [10], and (●) the GCM + ANN + PSO model proposed in this work.

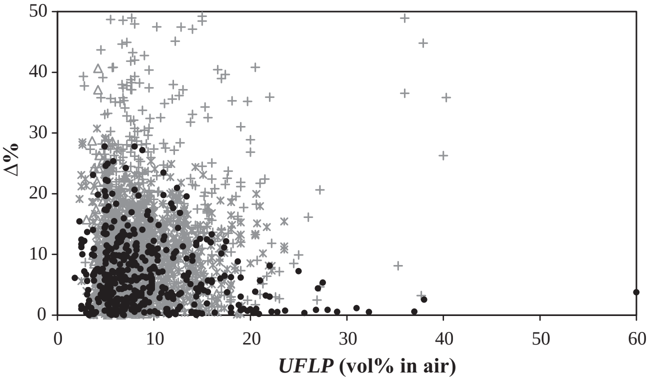

Fig. 7. Average absolute relative deviations found in the prediction of upper flammability limit percent in air using: (Δ) Pan et al. [7], (+) Pan et al. [11], (*) Gharagheizi [12], and (●) the GCM + ANN + PSO model proposed in this work.

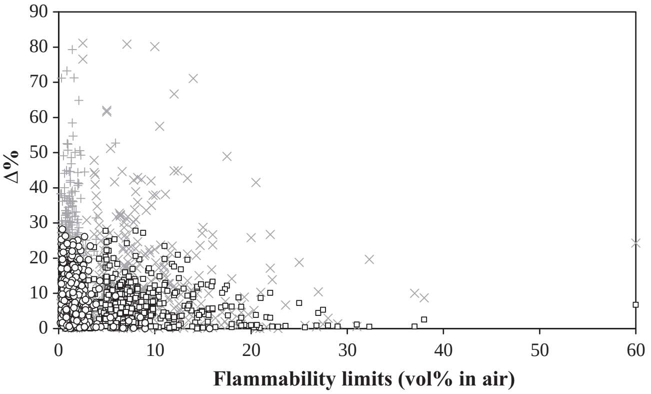

Fig. 8. Deviations found in the prediction of flammability limits percent in air using: artificial neural network with particle swarm optimization for *LFLP* (○) and *UFLP* (*), and artificial neural network with standard back-propagation for *LFLP* (+) and *UFLP* (×).

greater than 80%. Fig. 8 shows of the deviations found in the prediction of flammability limit temperatures (LFLT and ULFT) using the BPNN and the proposed model. The low deviations found with the proposed GCM + ANN + PSO model, indicate that can estimate *LFLP* and ULFP of organic compounds with better accuracy than other methods available in the literature. These results represent a tremendous increase in the accuracy to predict these important thermal properties.

5. Conclusions

In this work, the lower and upper flammability limits percentage in air of organic compounds have been estimated using a simple group contribution method implemented in an artificial neural network replacing standard back-propagation algorithm with an particle swarm algorithm.

Based on the results and discussion presented in this study, the following main conclusions are obtained: (i) The great differences in structure chemical and physical properties of the organic compounds considered in the study impose additional difficulties to the problem that the proposed method has been able to handle; (ii) The results show that the ANN + PSO can be properly trained and that the chosen architecture (42-6-2) can estimate the flammability limits percentage of organic compounds; (iii) The low deviations found with the proposed GCM + ANN + PSO model indicate that can estimate flammability limit percentage in air with better accuracy than others method available in the literature.

Acknowledgements

The author thank the Direction of Research of the University of La Serena (DIULS), and the Department of Physics of the University of La Serena (DFULS), by the special support that made possible the preparation of this paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tca.2010.09.018.

References

[1] M. Vidal, W.J. Rogers, J.C. Holste, M.S. Mannan, Process Safety Prog. 23 (2004) 47–55.
[2] L.G. Britton, K.L. Cashdollar, W. Fenlon, D. Furip, J. Going, B.K. Harrison, J. Niemeier, E.A. Ural, Process Safety Prog. 24 (2005) 12–28.
[3] D.E. Winterbone, Advanced Thermodynamics for Engineers, John Wiley & Sons, New York, 1997.
[4] G.W. Jones, Chem. Rev. 22 (1938) 1–26.
[5] E.A. Spakowski, Pressure Limit of Flame Propagation of Pure Hydrocarbons-Air Mixtures at Reduced Pressures, Res. Memo. E52H15, NACA, 1952.
[6] M. Hristova, Z. Zdravchev, S. Tchaouschev, J. Univ. Chem. Tech. Met. Sofia 37 (2002) 87–92.
[7] Y. Pan, J. Jiang, X. Ding, R. Wang, J. Jiang, AIChE J 56 (2010) 690–701.
[8] Y. Pan, J. Jiang, R. Wang, H. Cao, Y. Cui, J. Hazard. Mater. 168 (2009) 962–969.
[9] F. Gharagheizi, Energy Fuel 22 (2008) 3037–3039.
[10] F. Gharagheizi, J. Hazard. Mater. 170 (2009) 595–604.
[11] Y. Pan, J. Jiang, R. Wang, H. Cao, Y. Cui, Ind. Eng. Chem. Res. 48 (2009) 5064–5069.
[12] F. Gharagheizi, J. Hazard. Mater. 167 (2009) 507–510.
[13] J.A. Lazzús, Thermochim. Acta 489 (2009) 53–62.
[14] J.A. Lazzús, Int. J. Thermophys. 30 (2009) 833–909.
[15] J.A. Lazzús, Ind. Eng. Chem. Res. 48 (2009) 8760–8766.
[16] J.A. Lazzús, Fluid Phase Equilib. 289 (2010) 176–184.
[17] MatLab, The MathWorks Inc., Natick, 2002.
[18] T.E. Daubert, R.P. Danner, H.M. Sibul, C.C. Stebbins, Physical, Thermodynamic Properties of Pure Chemicals. Data Compilation, Taylor & Francis, London, 2000.