Revue Roumaine de Chimie, **2009**, *54*(9), 749–753

ACTUAL TRENDS IN THRESHOLD LIMIT VALUES EVALUATION: APPLICATION TO HALOGEN CONTAINING COMPOUNDS USING MULTILINEAR REGRESSION METHOD

Roxana DONE, Ioana STĂNCULESCU* and Cristina MANDRAVEL

Department of Physical Chemistry, Faculty of Chemistry, University of Bucharest, 4-12 Bd. Regina Elisabeta, District 3, Bucharest, 030018 Roumania, Tel.: +40213143508 room 2285; Fax +40213159249

Received January 29, 2009

In this paper, the multilinear regression method was applied for Threshold Limit Values (TLVs) evaluation. TLVs are occupational exposure values established by the American Conference of Governmental Industrial Hygienists (ACGIH) to prevent professional diseases. 10 experimental and calculated physical–chemical properties for a series of 17 halogenated alkanes were used. 154 equations, with 5 or 6 descriptors like: molecular weight (M), octanol-water partition coefficient (P), melting point (m. p.), boiling point, density, solvent accessible surface (So), van der Waals surface (SvdW), molecular volume (V), energies of highest occupied molecular orbital (EHOMO) and lowest unoccupied molecular orbital (ELUMO), with r^2 between 0.428-0.808 were obtained. It is obvious that better correlations are obtained when the number of descriptors increases. The best equation, selected after the highest values of the quality indices r^2 and F, contains 6 descriptors: logM, m.p., logP, ε_{HOMO}, ε_{LUMO}, logV and may be used to evaluate TLVs for halogen containing compounds.

INTRODUCTION

In a broad sense, any substance or mixture of substances having properties capable of producing adverse effects on the health or safety of a human is a hazardous chemical. After 1970 have been introduced specific regulations concerning the hazardous material classification. A chemical is considered perilous if it is determined to be cancer causing, toxic, corrosive, an irritant, a strong flammable or reactive and if it has a threshold limit value, assigned by the American Conference of Governmental Industrial Hygienists (ACGIH).[1,2]

TLVs are occupational exposure data, established to prevent professional diseases for the average person equating to a 150 pounds, male, age 25-44. TLVs have an important role in occupational medicine and industrial hygiene. They are expressed in mg/m^3 for particulate and in ppm for gaseous pollutants.[3]

Now ACGIH has proposed TLVs for almost 640 chemicals, which enter in the body by inhalation, eye, skin contact/absorption and accidentally ingestion[4] using experimental methods associated with epidemiological studies. Measurements of gas and vapor's concentrations may be realized by gas-chromatography (GC), total carbon analysis, FTIR spectroscopy, colorimetry, HPLC, specific sensors or classical methods of analysis and compared with TLVs.[5,6]

TLVs are generally reviewed and periodically updated.[7] There are different types of TLVs and assimilated defined concentrations as summarized in the third column of Table 1.

The task of evaluating all existing chemicals far exceeds the capacity of the toxicology professionals' world wide.[7] Many agencies (EPA-US Environmental Protection Agency, OSHA) use TLVs as a source for deriving ambient standards exposures by subjecting each TLV to a reduction factor.[4]

In literature, the correlation of TLVs with the occupational dust concentration on the development of chronic bronchitis was recently described.[8] It was assumed that there is no influence of the exposure dust concentration on the response under a certain limited value. Thus, the relationship is modeled in two segments. Under the threshold value the regression function is a constant, while it is increasing when the dust concentration exceeds the threshold. This is a special case of segmented regression.

* Corresponding author: ioana@gw-chimie.math.unibuc.ro

Table 1

Types of TLVs

No.	Types of TLVs	ACGIH notation	Other notations
1	**T**hreshold **L**imit **V**alue – **T**ime **W**eighted **A**verage (8 h work/day or 40 h/week exposure)	TLV– TWA	VL_8*, PEL/MAC**
2	**T**hreshold **L**imit **V**alue – **S**hort **T**erm **E**xposure **L**imit (max. 15 min. and 4 periods of exposure in a work day)	TLV –STEL	VL_s*
3	**T**hreshold **L**imit **V**alue – **C**eiling (this concentration may not be surpassed even 1 sec.)	TLV–C	-

* - Roumanian General Norms for Work Protection, adopted in 2002[3]
** - Permissible Exposure Limits (or Maximum Allowed Concentration) after the US Occupational Safety and Health Administration (OSHA).[7]

The correlation of TLVs for different pollutants with their fundamental physical-chemical properties [9] may be used for the TLV prediction for some compounds. Thus in the pioneering work of Filov and Liublina concerning the maximal admitted concentration (MAC) in the case of 12 halogenated hydrocarbons (carbon tetrachloride and o-dichlorobenzene were outliners) the following equation was found:

log MAC=3.6831 – 0.529log P with correlation coefficient r^2=0.8788 and squares residual sum Sr=0.7738.

There are differences between the prediction of TLVs and other properties (physical-chemical characteristics, toxicity and biological activity of substances).[10-20] These particularities are determined by the more restrained character of TLVs base data and its permanent updating process concomitantly with advancement of occupational medicine. Usually, TLVs are proposed by analysis of the monitoring occupational exposure data, clinical observations (lethal dose, indices of acute and chronic toxicity, etc.) and epidemiological research results.

In the present work correlations of TLVs with some experimental and calculated physical-chemical properties or reactivity indices for a series of halogen containing pollutants were searched using the multilinear regression method (MLR). This type of investigation may represent a useful tool in the control of risk associated with industrial emissions.[1]

RESULTS

A set of 17 halogen containing compounds were included in this study: methyl chloride (**1**), ethyl chloride (**2**), isopropyl chloride (**3**), dichloro methane (**4**), 1,1 dichloro ethane (**5**), 1,2 dichloro ethane (**6**), 1,2 dichloro propane (**7**), chloroform (**8**), 1,1,1 trichloro ethane (**9**), 1,2,3 trichloro propane (**10**), carbon tetrachloride (**11**), 1,1,1,2 tetrachloro ethane (**12**), 1,1,2,2 tetrachloro ethane (**13**), pentachloro ethane (**14**), hexachloro ethane (**15**), hexachloro cyclohexane (**16**), chloro bromo methane (**17**). Their physical-chemical data are from literature.[21-23] The TLVs[3] and the calculated properties for the studied class of compounds are given in Table 2.

Table 2

Calculated molecular properties for the halogenated alkanes

Compound	TLV-TWA mg/m^3	log P	ε_{HOMO} eV	ε_{LUMO} eV	S_{vdW} Å^2	S_G Å^2	V Å^3
1	75	0.88	-11.3381	1.5979	191.84	170.29	206.42
2	1000	1.22	-11.1535	1.4978	230.05	205.09	263.51
3	400	1.64	-11.1558	1.4074	268.47	232.55	314.44
4	174	1.15	-11.3897	0.5950	220.17	199.60	252.79
5	405	1.16	-11.4224	0.5823	257.21	229.81	306.21
6	30	1.59	-11.4165	0.6849	259.18	232.14	308.34
7	100	2.00	-11.2905	1.1173	293.84	253.27	355.67
8	10	1.61	-11.7710	-0.3031	246.36	225.08	296.75
9	1000	2.04	-11.9919	-0.2653	288.82	248.90	346.04
10	100	2.36	-11.3824	0.4514	316.44	276.11	397.37
11	30	3.37	-12.3785	-1.1165	275.12	244.77	337.94
12	20	2.42	-11.7928	-0.4845	312.81	268.33	385.04
13	20	2.06	-11.6553	-0.0738	307.17	270.23	384.38
14	40	2.65	-11.8694	-0.6808	336.76	287.28	421.57
15	5	3.25	-12.1822	-0.9669	364.51	300.53	454.21
16	0.3	4.65	-11.0318	-0.0391	374.58	363.65	590.10
17	700	1.03	-11.0868	0.1005	228.41	210.61	272.38

Continuing our precedent work[22] in this paper we present the calculation results for equations with five (Table 3) and six (Table 4) descriptors.

Table 3

The best equations with 5 descriptors

No.	Descriptors (X_i)	r^2	F	Coefficients (a_i)	Compounds
1	logM, m.p., logP, ε_{HOMO}, S_{vdW}	0.785	7.291	-28.438; -6.251; -4.345E-3; -0.244; -0.914; 13.498	**1-16**
2	logM, m.p., logP, ε_{HOMO}, S_G	0.717	7.916	-53.608; -9.764; -3.544E-3; -0.487; -1.729; 23.692	**1-16**
3	logM, m.p., logP, ε_{HOMO}, logV	0.805	8.271	-40.833; -9.527; -3.768E-3; -0.552; -1.727; 17.199	**1-16**
4	logM, m.p., logP, ε_{LUMO}, S_{vdW}	0.788	7.443	-20.27; -9.975; -3.694E-3; -0.228; -0.768; 17.751	**1-16**
5	logM, m.p., logP, ε_{LUMO}, S_G	0.750	5.999	-35.534; -15.415; -1.989E-3; -0.564; -1.366; 29.691	**1-16**
6	logM, m.p, logP, ε_{LUMO}, lgV	0.778	7.028	-22.217; -16.984; -2.09E-3; -0.676; -1.544; 24.194	**1-16**
7	logM, m.p., logP, S_{vdW}, logV	0.743	5.771	-14.412; -2.932; -4.591E-3; -0.039; 23.527; -13.969	**1-16**
8	logM, m.p, ε_{HOMO}, S_{vdW}, lgV	0.772	6.754	-28.154; -7.084; -5.098E-3; -0.974; 13.939; -0.337	**1-16**
9	logM, m.p., ε_{HOMO}, S_G, logV	0.749	5.968	-50.963; -10.325; -4.924E-3; -1.709; 28.849; -5.79	**1-16**
10	logM, m.p., ε_{LUMO}, S_{vdW}, logV	0.777	6.969	-19.387; -10.579; -4.369E-3; -0.779; 19.044; -1.3	**1-16**

Table 4

The best equations with 6 descriptors

No.	Descriptors (X_i)	r^2	F	Coefficients (a_i)	Compounds
1	logM, m.p., d, logP, S_{vdW}, logV	0.702	3.137	52.557; 13.906; 6.542E-3; -6.696; -0.437; -62.684; 31.816	**1-15**
2	logM, m.p., logP, ε_{HOMO}, ε_{LUMO}, S_{vdW}	0.789	5.606	-22.837; -8.971; -3.89E-3; -0.23; -0.273; -.552; 16.626	**1-16**
3	logM, m.p., logP, ε_{HOMO}, ε_{LUMO}, S_G	0.798	5.941	-53.374; -10.167; -3.466E-3; -0.491; -1.665; -0.07; 24.268	**1-16**
4	logM, m.p., logP, ε_{HOMO}, ε_{LUMO}, logV	0.808	6.323	-37.916; -11.957; -3.336E-3; -0.589; -1.358; -0.415; 19.821	**1-16**
5	logM, m.p., logP, ε_{LUMO}, S_{vdW}, logV	0.796	5.839	-22.648; -13.561; -2.997E-3; -0.413; -1.142; 11.269; 10.371	**1-16**
6	logM, m.p., ε_{HOMO}, ε_{LUMO} , S_{vdW}, logV	0.777	5.236	-21.466; -10.162; -4.509E-3; -0.207; -0.647; 17.745; -0.53	**1-16**
7	logM, m.p., ε_{HOMO}, ε_{LUMO}, S_G, logV	0.751	4.533	-58.42; -8.234; -5.096E-3; -2.063; 0.378; 35.79; -12.791	**1-16**
8	logM, d, logP, ε_{LUMO}, S_G, logV	0.707	3.213	30.931; -5.014; -2.635; -0.645; -1.028; -70.231; 61.059	**1-15**
9	logM, logP, ε_{HOMO}, ε_{LUMO}, S_{vdW}, logV	0.783	5.406	-38.084; -15.431; -0.845; -1.31; -0.746; -5.258; 28.363	**1-16**
10	logM, logP, ε_{HOMO}, ε_{LUMO}, S_G, logV	0.787	5.551	-8.384; -16.712; -0.876; -0.516; -1.263; -38.513; 52.7	**1-16**
11	b.p., d, logP, ε_{LUMO}, S_G, logV	0.700	3.106	29.905; -3.248E-3; -2.883; -0.552; -0.598; -65.549; 53.038	**1-15**
12	m.p., d, logP, ε_{LUMO}, S_{vdW}, logV	0.701	3.124	-7.624; 2.041E-3; -3.774; -0.411; -0.757; 20.196; -13.304	**1-15**
13	m.p., d, logP, ε_{LUMO}, S_G, logV	0.726	3.526	38.694; 3.476E-3; -4.323; -0.696; -0.916; -70.932; 55.524	**1-15**
14	d, logP, ε_{HOMO}, ε_{LUMO}, S_G, logV	0.705	3.182	16.864; -3.172; -0.651; -0.796; -0.343; -55.698; 45.325	**1-15**

In these tables are listed only equations where correlation coefficient r^2 is higher than 0.7 from the total of 84 for 5, respectively 70 for 6 descriptors obtained equations. The *F* factor is a measure of the regression relationship being calculated by the

formula: $F = \frac{MSR}{MSE}$, where MSR = mean square regression and MSE = mean square error.[24]

DISCUSSION

Different combinations of 5 and 6 descriptors from the total of 10 considered in this study (4 experimental and 6 calculated) were tried.

The equations used were of the type:

$$Y = a_0 + \sum_{i=1}^{5/6} a_i X_i$$

where Y = log TLV, a_0 = free term; a_i – coefficients (i = 1,…, 5/6), X_i – descriptors (physical-chemical experimental or calculated data). Quality of these equations may be appreciated after the values of r^2 and F factor shown in the columns 3 and 4 of Tables 3 and 4.

From Table 3 analysis one observes that the highest values of the quality indices r^2 (0.805) and F (8.271) correspond to correlations including log M, m.p., log$P$, ε_{HOMO} and log V as descriptors (2 experimental and 3 calculated). In Table 4 the highest values of quality indices (r^2=0.808, F=6.333) appear for correlations including log$M$, m.p., log$P$, ε_{HOMO}, ε_{LUMO} and log V descriptors (2 experimental and 4 calculated). These show that a combination of experimental and theoretical data may provide a better estimation of TLVs.

In Table 4, equations 1, 8, 11, 12, 13, 14, respectively 2, 3, 4, 5, 6, 7, 9 and 10, differ by the number of the compounds considered for regression. Comparative analysis of the data comprised in the Tables 3 and 4 shows that the equations quality improves with the number of descriptors considered. Also, the comparison with our previous results[22] on the TLVs estimation shows that using 3 or 4 descriptors gives poorer quality equations. The correlations with 5 respectively 6 descriptors were distributed after the r^2 values as follows in Table 5.

For future research the opportunity to realize calculations including higher combinations of experimental and calculated molecular descriptors and also extension of the studied halogenated compounds group will be considered.

Table 5

Number of relations with 5, respectively 6 descriptors

r^2 values interval	No. of relations with 5 descriptors	No. of relations with 6 descriptors
0.4-0.499	3	1
0.5-0.599	17	8
0.6-0.699	54	47
0.7-0.799	9	13
0.8-0.899	1	1

CALCULATION DETAILS AND SELECTION OF THE DATA

TLV's, respectively, MAC's or VL_8 values of exposure limits are from literature sources. [3,7] The values of the physical properties: melting points – m.p., boiling points – b.p., densities – d are also based on literature data. [21-23] Molecular weight – M is considered a special type of descriptor because the toxicity of a chemical increases with the complexity of the structure.[3]

The values of partition coefficient P between octanol/and water phases were calculated after the incremental method of Ghose, Pritwchett and Crippen, implemented in the HYPERCHEM program.[25,26] Calculated log P values were compared with those experimentally determined, comprised in EPI Suite v. 3.20, developed by the EPA's Office of Pollution Prevention Toxics and Syracuse Research Corporation and small differences were found (between 0.02 and 0.89).[27] The other molecular properties of this class of compounds: grid (solvent accessible) surface – S_G, approximate (van der Waals) surface – S_{vdW}, molecular volume – V, energies of frontier orbitals HOMO (**H**ighest **O**ccupied **M**olecular **O**rbital) – ε_{HOMO}, LUMO (**L**owest **U**noccupied **M**olecular **O**rbital) –ε_{LUMO} were calculated for the structures optimized with the AM1 method in the RHF approximation, *in vacuo* with the HYPERCHEM program.[26] Minimum energy structures were obtained by optimization with the Polack-Ribiere algorithm with a RMS gradient of 0.01 kcal/mol·Å.

Multilinear regression (**MLR**) calculations in the frame of MATHCAD 7 professional program were done.[24]

CONCLUSIONS

Correlation of TLV's for 17 halogenated alkanes with molecular properties using multilinear regression was done. The most efficient combinations of 5, respectively 6 descriptors from the group of 10, considering the values of quality indices indicated the best 10, respectively 14, determined equations.

A value of r^2 higher than 0.7 means a moderate prediction. The quality of the equations indicates that the theoretical estimation of TLVs for other related compounds may be a good alternative to experiment.

Acknowledgement. Authors are grateful to Professor Gheorghe Surpățeanu from the University of Dunkerque, France and to Professor Cezar Bendic from the University of Bucharest, for the kind help provided with the calculation facilities and data analysis.

REFERENCES

1. R. Niesink, M. A. Hollinger and J. De Vries, "Toxicology: Principles and Appllications", CRC Press, Boca Raton, 1996, p. 1141.
2. C. E. Gorman (Ed.), Working safely with chemicals in the laboratory, 2nd edition, GP Corporation, Schenectady, 1995, p. 105-123.
3. I. Silion and C. Corduneanu, "Bazele medicinii muncii", Editia a 3-a, Ed. PIM, Iasi, 2003, p. 316.
4. V. Chiosa, I. Stănculescu and C. Mandravel, "Evaluarea toxicitatii poluantilor atmosferici", Ed. Univ. Buc., 2007, p. 140.
5. C. Mandravel and R. Stanescu-Dumitru, "Metode fizico-chimice aplicate la masurarea noxelor in mediul profesional", Ed. Acad. Rom., 2003, p. 259-266, 270-273.
6. C. C. Chang, R. H. Shic and T. Y. Chang, *Int. Archiv. Occup. Env. Health*, **2006**, *79*,135.
7. J. S. Smith and J. M. Mendeloff, *Risk Analysis*, **1999**, *19*, 1223-1234.
8. H. Kuchenhoff and R. J. Carroll, *Statistics in Medicine*, **1997**, *16*, 169-188.
9. V. A. Filov and E. A. Liublina, *Biofizika*, **1962**, *7*, 1-8.
10. A. Beteringhe and A. T. Balaban, *ARKIVOC*, **2004**, *1*,163-182.
11. A. E. Soffers, M. G. Boersma, H. J. Waes, J. Vervoort, B. Tyrakovska, J. Hermens and I. M. Rietjins, *Toxicology in Vitro,* **2001**, *15*, 539-551.
12. S. D. Bolboaca and L. J. Jantschi, *Terap. Farmacol. Toxicol. Clin.*, **2006**, *X*, 113-117.
13. S. Arulmozhiraya and M. Morita, *Chem. Res. Toxicol.*, **2004**, *17*, 348-356.
14. S.Trohalaki, E. Gifford and R. Pachter, *Comput. Chem.*, **2000**, *24*, 421.
15. D.A. Whaley, M. D. Attfield, E. J. Bedillion and K. M. Walter, *Ann. Occup. Hyg.*, **2000**, *44*, 361-374.
16. L.Tarko and O.Ivanciuc, *MATCDY*, **2001**, *44*, 201-214.
17. O. Ivanciuc, T. Ivanciuc and A. T. Balaban, *Tetrahedron*, **1998**, *54*, 9129-9142.
18. A. R. Katritzky, M. Lan, V. S. Lobanov and M. Karelson , *J. Phys. Chem.*, **1996**, *100*, 10400.
19. O. Ivanciuc, *Internet Electronic J. Mol. Design*, **2002**, 1, 252-268.
20. M. Karelson and A. Perkson, *Computers & Chemistry*, **1999**, *23*, 49-59.
21. *** "Ullmann's Encyclopedia of industrial chemistry", 5th edition, Wiley –VCH-69451, Wennheim, Germany, 1997.
22. I. Stănculescu, R. Done and C. Mandravel, Proc. *32rd Amer. Rom. Acad. Congr*., Boston, USA, **2008**, July 22-24, 615-618.
23. www.chemfinder.cambridgesoft.com
24. Mathcad7 Professional program, 1986-1997, MathSoft, Inc.
25. A. K. Ghose and G. M. Crippen, *Chem. Inf. Comp. Sci.*, **1987**, *27*, 21.
26. HYPERCHEM program, Release 6.01 for Windows, Hypercube Inc. 2000.
27. R. Done, G. Mîndrilă and I. Stănculescu, I., *Anal. Univ. Buc. Chimie*, 2007, XVI, II, 45-49.