Using fuzzy artificial neural network TSK (Takagi, Sugeno, Kang) for approximation and prediction of dissociation energy of C-X-bonds (X=F, Cl, Br, I) in halogenated hydrocarbons

VLADIMIR E. TUMANOV, ELENA S. AMOSOVA, BULAT N. GAIFULLIN, ANDREY I. PROKHOROV
Department of Computing and Information Resources
Institute of Problems of Chemical Physics RAS
142432, Chernogolovka, Semenov ave, 1
RUSSIAN FEDERATION
tve@icp.ac.ru

Abstract: The Takagi–Sugeno–Kang *(*TSK*)* fuzzy artificial neural network has been used for approximation of dissociation energies of C-X-bonds (X=F, Cl, Br, I) in halogenated hydrocarbons by the experimental data. Characteristics of molecule: electronegativity, force constant of the bond, the atom size of halogen served as variables. The comparison of predictions by the developed fuzzy network with the experimental data on the test sample is given. The obtained results are in good agreement with the experimental data.

Key-Words: fuzzy neural network, Takagi-Sugeno-Kang model, bond dissociation energy, halogenated hydrocarbon, electronegativity, bond force constant, halogen atom size.

1 Introduction

Mathematical apparatus of artificial neural networks (ANN) and fuzzy knowledge bases is currently widely used in applied problems of automated processing of the scientific data in interdisciplinary research in chemistry [1-3]. One of such topical applied problems is approximation and prediction of bond dissociation energy of organic molecules by the experimental data.

Papers [4-5] describe using the ANN technology for prediction of bond dissociation energies (by the experimental data and with chemical descriptor set).

However at present there are a few papers on using fuzzy logic in chemical kinetics and thermochemistry (see the review in [6]). Some papers on estimation of bond dissociation energies using fuzzy knowledge base and fuzzy neural networks [7] on the data base samples [8] are worth mentioning.

The purpose of this paper is using fuzzy neural network for approximation and prediction of dissociation energy of C-X-bonds (X=F, Cl, Br, I) by the experimental characteristics of the molecule and parameters of the chemical bond.

2 Problem Formulation

In [9] an array of dissociation energies of C-X-bonds (X=F, Cl, Br, I) has been calculated and regression models have been developed

$$D_{C-X} = \omega_1 \sqrt{D_{X-X} D_{C-C}} + \omega_2 br_{XX} + \omega_3 ,$$

where $D_{\text{F-F}}$ = 158.670 ± 0.096 kJ/mol, $D_{\text{Cl-Cl}}$ = 242.58 ± 0.004 kJ/mol, $D_{\text{Br-Br}}$ = 193.859 ± 0.120 kJ/mol, $D_{\text{I-I}}$ = 152.25 ± 0.57 kJ/mol, $D_{\text{C-C}}$ is calculated according the enthalpies of formation of free radicals.

Regression models have been developed for the following type of compounds: RCH_2-X, RCH-X, *cyclo*-[$(CH_2)_n$CH-X], RPh-X, R=CH-X, $RCCH_2$-X, RC(O)-X, PhC(O)-X, where R and Ph are alkyl and phenyl radical correspondingly.

The developed regression models allow one to form fuzzy knowledge base and to use the fuzzy neural network of Takagi-Sugeno-Kang (TSK) for approximation and prediction of dissociation energies of C-X-bonds by the experimental and calculated data for compound classes mentioned above.

3 Problem Solution

In given research the generalized inference scheme of the TSK model using M=9 rules and N=3 variables x_j can be presented as follows [10]:

Base of rules

R_1: if x_1 is $A_1^{(1)}$ and x_2 is $A_2^{(1)}$ and x_3 is $A_3^{(1)}$ then

$$y_1 = p_{01} + \sum_1^N p_{1j} x_j$$

…

R_M: if x_1 is $A_1^{(M)}$ and x_2 is $A_2^{(M)}$ and x_3 is $A_3^{(M)}$ then

$$y_M = p_{M0} + \sum_1^N p_{Mj} x_j$$

where $A_i^{(k)}$ is the value of linguistic variable x_i for R_i rule with the membership function

$$\mu_A^{(k)}(x_i) = \frac{1}{1+\left(\frac{x_i - c_i^{(k)}}{\sigma_i^{(k)}}\right)^{2b_i^{(k)}}}, i = \overline{1,N}, \; k = \overline{1,M}$$

In the TSK fuzzy neural network the intersection of R_k rules is determined by the membership function as multiplication, i.e.

$$\mu_A^{(k)}(x) = \prod_{j=1}^N \frac{1}{1+\left(\frac{x_j - c_j^{(k)}}{\sigma_j^{(k)}}\right)^{2b_j^{(k)}}}$$

By M rules of detachment the composition of initial results of the network is determined by the following formula:

$$y(x) = \frac{\sum_{k=1}^M \omega_k y_k(x)}{\sum_{k=1}^M \omega_k}$$

where $y_k(x) = p_{k0} + \sum_{j=1}^N p_{kj} x_j$, $\omega_k = \mu_A^{(k)}(x)$

The general expression for functional dependence in the TSK network is given as:

$$y(x) = \frac{1}{\sum_{k=1}^M \prod_{j=1}^M \mu_A^k(x_j)} \sum_{k=1}^M (p_{k0} + \sum_{j=1}^N p_{kj} x_j) \prod_{j=1}^N \mu_A^{(k)}(x_j)$$

The TSK fuzzy neural network is described by multilayer structural network of 5 layers where the first and the third layers are parametric.

The learning is supervised (for every input vector there is a desired output value). The set is divided into two parts: the learning and the validating one.

The first variable is the linguistic one describing the structural group of alkyl halogens, the second one is the molecule electronegativity given as $\sqrt{D_{X-X} D_{C-C}}$, the third variable is the product of multiplication of force constant of the chemical bond $br_{X\text{-}X}$ by the atom size of X.

It should be noted that the first variable is purely linguistic. Numerical values for the value set of this linguistic variable have been chosen by random to secure the possibility of training of the first network layer by backpropagation method. Inversion of the pseudoinverse matrix has been made by the iterative method of Ben-Israel [11].

The set of the experimental data of dissociation energy of C-X bonds includes 195 values, of which 20 are the test sample.

Table 1. Learning results

Compound	D_{calc}	D_{exp} [9]	ΔD
	kJ/mole		
CH_3CH_2-Br	295.0	294.5±2.1	-0.5
CH_3CH-{Cl}CH_3	356.6	355.3±7.0	0.3
CH_2-{F}$(CH_2)_3CH_3$	463.1	461.9±2.9	1.2
CH_3CH-{I}$(CH_2)_3CH_3$	233.1	232.4±2.8	0.7
cyclo-[CH-{Cl}$(CH_2)_4$]	346.7	346.7±4.2	0.0
cyclo-[CH-{Br}$(CH_2)_5$]	292.0	291.3±2.8	0.7
$(CH_3)_2CHCH_2$-{F}	461.3	460.7±2.3	0.6
$(CH_3)_3C$-{Cl}	350.1	349.2±3.0	0.9
CH_2=CH-{I}	275.2	276.1±5.0	-0.9
CH_2=$CHCH_2$-{Br}	236.0	235.2±3.1	0.8
HC≡C-{F}	432.2	430.8±0.3	1.4
C_6H_5-{I}	281.3	280.8±13.9	0.5
$C_6H_5CH_2$-{Cl}	310.0	309.3±4.1	0.7
4-$CH_3C_6H_4$-{Cl}	410.2	411.4±2.6	-1.2
CH_2-{I} CH_2OH	236.3	234.8±7.0	1.5
CH_3C-{F}(O)	511,4	512.4±5.0	-1.0
C_6H_5C-{Br}(O)	254.0	253.6±11.6	0.4
$CH_3C(O)CH_2$-{F}	440.7	441.0±1.9	-0.3
CH_3OCH_2-{I}	238.9	239.6±1.5	-0.7
{Cl}-C≡N	420.0	418.5	1.5

The network learning has been completed by 15000 iterations. In Table 1 the results of prediction on the test sample are given. As seen from the Table 1, there is a good agreement between the values predicted by the fuzzy neural network and the experimental data. The mean-square prediction error for $\Delta D = D_{calc} - D_{exp}$ on the test sample is 0.33 ± 0.85 kJ/mol.

4 Conclusion

For the first time the Takagi-Sugeno-Kang fuzzy neural network has been used for approximation of dissociation energy of C-X bonds (X= F, Cl, Br, I)

Prediction of bond dissociation energies within the TSK model gives satisfying results.

Evaluation of the $D_{C\text{-}F}$ value for $CH_3C(O)CH_2F$, according to the regression model [9] is 500.7 kJ/mol, while the developed network predicts 501.1 kJ/mol, which is in a good accordance.

The fuzzy neural networks described in this paper according to the TSK model allows one to predict dissociation energies of C-X bonds in halogenated hydrocarbons with a sufficient degree of accuracy. This allows one to solve a number of topical problems in determining the reactivity of given compounds in atmospheric chemistry.

References:

[1] X1. Baskin I.I., Palyulin V.A., Zefirov N.S. Application of Artificial Neural Networks in Chemical and Biochemical Studies. *Vestnik Moskovskogo Universiteta, Ser. 2. Khim.* — 1999. Vol. 40(5). — P. 323-326 (Russ.).

[2] X2. Himmelblau D.M., Applications of Artificial Neural Networks in Chemical Engineering, *Korean Z Chem. Eng.*, Vol. 17, N. 4, 2000, pp. 373-392.

[3] X3. Gasteiger J., Zupan J., Neural networks in chemistry, *Angev. Chem. Int. Ed. Engl.*, Vol. 32, 1993, pp. 503-527.

[4] X4. Lazarev D. Yu., Tumanov V.E., Using of Artificial Neural Network to Assess the Bond Dissociation Energies of Organic Molecules, *INORGANIC MATERIALS English translation of selected articles from Zavodskaya Laboratoriya. Diagnostika Materialov*, Vol. 80, N. 5, 2014, pp. 74-76.

[5] X5. Urata S., Takada A., Uchimaru T., Chandra A.K., Sekiya A., Artificial neural network study for the estimation of the C-H bond dissociation enthalpies, *Journal of Fluorine Chemisty*, Vol. 116, 2002, pp. 163-171.

[6] X6. Tumanov V.E. Application of the Artificial Neural Networks and Fuzzy Logic for the Prediction of Reactivity of Molecules in Radical Reactions // *In: Computational Problems in Engineering. Series: Lecture Notes in Electrical Engineering* (Ed. N. Mastorakis, V. Mladenov.) — Switzerland: Springer International Publishing, Vol. 307, 2014, pp. 261-269.

[7] X7. Lazarev D.Yu., Tumanov V.E., Prediction of bond dissociation energies of organic molecules by kinetic data of the radical reactions using the fuzzy knowledge base, *BIT,s 3rd Annual Conference and EXPO of AnalytiX-2014*, (April 25-28 2014. Dalian, China). 2014, p. 379.

[8] X8. Tumanov V., Gaifullin G., Subject-oriented science intelligent system on physical chemistry of radical reactions, *Modern Advances in Intelligent Systems and Tools*, Vol. 431, 2012, pp. 121-126.

[9] X9. Tumanov V.E., Bond dissociation energy of C-X (X= F, Cl, Br, I) bonds in halogen-containing hydrocarbons: correlation ratios with electronegativity, force constants of bonds and radii of atom X, *Butlerov Communications*, Vol.40, N.8, 2015, pp. 140-146.

[10] X10. Jang J., Sun C., Mizutani E., *Neuro-fuzzy and Soft Computing*, and, Prentice Hall 1997.

[11] X11. Shinozaki N., Sibuya M., Tanabe K, Numerical algorithms for the Moore-Penrose inverse of a matrix: interative methods, *Annals of the Institute of Statistical Mathematics*, Vol. 24, N. 1,1972, pp. 621-629.