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

$$D_{C-X} = \omega_1 \sqrt{D_{X-X}} + D_{C-C} + \omega_2 r_{xx} + \omega_3,$$

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

Regression models have been developed for the following type of compounds: RCH2-X, RCH-X, cyclo-[{(CH2)nCH-X}], RPh-X, R=CH-X, RCCH2-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 \( x_i \) can be presented as follows [10]:

Base of rules

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

\[
y_1 = p_{0i} + \sum_{j=1}^{N} p_{1j} x_j
\]

\[
... 
\]

\[ y_M = p_{M0} + \sum_{j=1}^{N} p_{Mj} x_j \]

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

\[
\mu_{A_i}^{(k)}(x_i) = \frac{1}{1 + \left( \frac{x_i - c_i^{(k)}}{\sigma_i^{(k)}} \right)^{2\beta_i^{(k)}}}, \quad i = 1, N, \quad k = 1, M
\]

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

\[
\mu_{A_i}^{(k)}(x) = \prod_{j=1}^{N} \frac{1}{1 + \left( \frac{x_j - c_i^{(k)}}{\sigma_j^{(k)}} \right)^{2\beta_j^{(k)}}}
\]

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

\[
y(x) = \sum_{k=1}^{M} \omega_k y_k(x)
\]

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

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

\[
y(x) = \frac{1}{\sum_{k=1}^{M} \prod_{j=1}^{N} \mu_{A_i}^{(k)}(x_j)} \sum_{k=1}^{M} \left( p_{k0} + \sum_{j=1}^{N} p_{kj} x_j \right) \prod_{j=1}^{N} \mu_{A_i}^{(k)}(x_j)
\]

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 \( br_{X,X} \) 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

<table>
<thead>
<tr>
<th>Compound</th>
<th>( D_{calc} ) KJ/mole</th>
<th>( D_{exp} ) [9] KJ/mole</th>
<th>( \Delta D ) KJ/mole</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH(_2)CH(_2)-Br</td>
<td>295.0</td>
<td>294.5±2.1</td>
<td>-0.5</td>
</tr>
<tr>
<td>CH(_2)-{Cl}CH(_3)</td>
<td>356.6</td>
<td>355.3±7.0</td>
<td>0.3</td>
</tr>
<tr>
<td>CH(_2)-{F}(CH(_2))CH(_3)</td>
<td>463.1</td>
<td>461.9±2.9</td>
<td>1.2</td>
</tr>
<tr>
<td>CH(_3)-{Cl}(CH(_2))(_3)</td>
<td>233.1</td>
<td>232.4±2.8</td>
<td>0.7</td>
</tr>
<tr>
<td>cyclo-[CH-{Cl}(CH(_2))(_3)]</td>
<td>346.7</td>
<td>346.7±4.2</td>
<td>0.0</td>
</tr>
<tr>
<td>cyclo-[CH-{Br}(CH(_2))(_2)]</td>
<td>292.0</td>
<td>291.3±2.8</td>
<td>0.7</td>
</tr>
<tr>
<td>(CH(_2))CH(_2)-{F}</td>
<td>461.3</td>
<td>460.7±2.3</td>
<td>0.6</td>
</tr>
<tr>
<td>(CH(_2))C-{Cl}</td>
<td>350.1</td>
<td>349.2±3.0</td>
<td>0.9</td>
</tr>
<tr>
<td>CH(_2)-CH-{I}</td>
<td>275.2</td>
<td>276.1±5.0</td>
<td>-0.9</td>
</tr>
<tr>
<td>CH(_2)-CH(_2)-{Br}</td>
<td>236.0</td>
<td>235.2±3.1</td>
<td>0.8</td>
</tr>
<tr>
<td>HC={C-{F}</td>
<td>432.2</td>
<td>430.8±0.3</td>
<td>1.4</td>
</tr>
<tr>
<td>C(_2)H(_2)-{I}</td>
<td>281.3</td>
<td>280.8±13.9</td>
<td>0.5</td>
</tr>
<tr>
<td>C(_2)H(_2)-{Cl}</td>
<td>310.0</td>
<td>309.3±4.1</td>
<td>0.7</td>
</tr>
<tr>
<td>4-CH(_3)C(_6)H(_5)-{Cl}</td>
<td>410.2</td>
<td>411.4±2.6</td>
<td>-1.2</td>
</tr>
<tr>
<td>CH(_2)-{I} CH(_2)OH</td>
<td>236.3</td>
<td>234.8±7.0</td>
<td>1.5</td>
</tr>
<tr>
<td>CH(_2)-{F}(CH(_2))(_3)</td>
<td>511.4</td>
<td>512.4±5.0</td>
<td>-1.0</td>
</tr>
<tr>
<td>C(_2)H(_2)-{Br}(O)</td>
<td>254.0</td>
<td>253.6±11.6</td>
<td>0.4</td>
</tr>
<tr>
<td>CH(_2)(O)CH(_2)-{F}</td>
<td>440.7</td>
<td>441.0±1.9</td>
<td>-0.3</td>
</tr>
<tr>
<td>CH(_2)OCH(_2)-{I}</td>
<td>238.9</td>
<td>239.6±1.5</td>
<td>-0.7</td>
</tr>
<tr>
<td>{Cl}-C(_3)=N</td>
<td>420.0</td>
<td>418.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

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 satisfactory results.

Evaluation of the $D_{C,F}$ value for CH$_3$C(O)CH$_2$F, 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:


