Evaluation of the rate constants reactions of the alkyl, allyl and aryl radicals with hydrocarbons with use the artificial of neural network

Tumanov Vladimir Evgen'rvich

Abstract:

This paper discusses the use of feed-forward artificial neural network to predict the rate constants of organic molecules in the bimolecular radical reactions $R + R_1H$ in the liquid phase on the experimental data. The hybrid algorithm of calculation of rate constants of bimolecular radical reactions on the experimental thermochemical data and an empirical index of the reactionary center is offered. This algorithm uses an artificial neural network for a prediction of a classical barrier of bimolecular radical reaction at a temperature of 298 K, a database of experimental characteristics of reaction and Arrhenius's formula for calculation of a rate constant. Results of training and a prediction of the network are discussed. Results of comparison of logarithms of the calculated and experimental of the rate constants are given.

Keywords:

artificial neural network, algorithm, empirical index of the reactionary center, rate constants, free radical reaction, reactivity of organic molecules.

Оценка констант скорости реакций алкильных, аллильных и арильных радикалов с углеводородами искусственной нейронной сетью

Туманов Владимир Евгеньевич

Аннотация:

В статье обсуждается применения искусственных нейронных сетей прямого распространения для предсказания констант скорости органических молекул в реакциях $R + R_1H$ в жидкой фазе. Предложен гибридный алгоритм вычисления констант скорости радикальных бимолекулярных реакций отрыва по экспериментальным термохимическим данным и эмпирическому индексу реакционного центра реакций. Этот алгоритм использует искусственную нейронную сеть для предсказания классического потенциального барьера бимолекулярной радикальной реакции при температуре 298° К, базу данных экспериментальных характеристик реакции и формулу Аррениуса для вычисления констант скорости. Обсуждаются результаты обучения нейронной сети. Проведено сравнение логарифмов вычисленных и экспериментальных констант скорости.

Ключевые слова:

искусственные нейронные сети, алгоритм, эмпирический индекс

реакционного центра, константы скорости, радикальная реакция, реакционная способность органических молекул.

Introduction

Currently artificial neural network (ANN) is widely used in solving applied problems of automated processing of scientific data. The main fields of ANN application in chemical, biochemical and informatics studies are given in [1-5]. Most works in this area are devoted to the correlation between the structure of chemical compounds and the physicochemical properties or biological activity they showed. In physical chemistry, the main directions of ANN application are simulation of chemical processes and simulation of dynamic properties of molecules and the systems.

One of the actual tasks is prediction of reactivity of molecules in chemical reactions (activation energy and rate constant). A number of research teams engaged in the study of prediction of reactivity of organic molecules in radical reactions.

Prediction of reactivity was done using various approaches. Using the general regression neural network, quantum chemical descriptors, functional density theory (DFT) taking as a basis a ratio of «structure-property» (quantitative structure-activity relationship) [6]; using multi-layer perception, chemical descriptors, experimental data taking as a basis a ratio of "structure-property» [7]; using regular feedforward neural network with a backpropagation training algorithm, experimental data described by kinetic differential equations [8,9]; using feedforward neural network, kinetic curve, without a kinetic model [10]; using feedforward neural network with a backpropagation training algorithm, experimental thermochemical and kinetic data [11].

In physical chemistry of radical reactions the large amount of experimental data on reactivity (specific reaction rate or activation energies) of molecules in radical reactions in the liquid phase was accumulated [12]. Knowledge of reactivity of organic molecules in radical reactions is necessary for the development of new organic materials, design of new drugs, design of technological processes, planning and conducting of scientific experiment, training of students and graduate students. Therefore, the development of ANN based on existing experimental data to predict reactivity of organic molecules in radical reactions is important task.

This paper discusses the use of feed-forward artificial neural network to predict reactivity of organic molecules in bimolecular radical reactions $R + R_1H$ in the liquid phase on experimental data. The real work is continuation of researches [11].

Problem Formulation

Experimentally, the activation energy (E) or a classical potential barrier (E_e) determines reactivity of organic molecules in a radical reaction:

$$E_e = E + 0.5(hL\nu_i - RT) \tag{1}$$

where: ν_i is the frequency of the stretching vibrations for the bond being broken, R is the gas constant, h is the Planck constant, L is the Avogadro number, and T is the reaction temperature (K).

Specific rate constant (k) of chemical reaction is calculated by the formula:

$$k = nA_0 \exp(-E/RT) \tag{2}$$

where: A_0 is the collision frequency per one equireactive bond, n is the number of equireactive bonds in a molecule.

When designing the information space for ANN predictions of the reactivity the functional relationship between the reactivity of the chemical reaction and the thermochemical properties (enthalpy of reaction - ΔH) is used.

N.N. Semenov was the first to pay attention to functional relationship between reactivity and reaction enthalpy (known as Polanyi – Semenov's ratio [13]):

$$E = B - \gamma\Delta H \tag{3}$$

where B and γ are the empirical coefficients.

The works [14, 15] proposed empirical models of elementary radical reactions, which allowed constructing non-linear correlation dependences between the classical potential barrier of radical reaction and its thermochemical properties:

- approximation of the above mentioned dependence in the work [14] by the parabola:

$$br_e = \alpha\sqrt{E_e - \Delta H_e} - \sqrt{E_e} \tag{4}$$

- approximation of the above mentioned dependence in the work [15] in the form of the tacitly set curve:

$$br_e = D_{ei}^{1/2}\ln\left(\frac{D_{ei}^{1/2}}{D_{ei}^{1/2} - E_e^{1/2}}\right) + \alpha D_{ef}^{1/2}\ln\left(\frac{D_{ef}^{1/2}}{D_{ef}^{1/2} - (E_e - \Delta H_t)^{1/2}}\right) \tag{5}$$

Under the proposed empirical models assuming the harmonic stretching vibrations, the reaction of the radical abstraction R°· + R₁H → RH +R°₁ (where R°·and R°₁ are alkyl radicals, and RH and R₁H are hydrocarbon molecules) has the following parameters [14,15]:

1. Enthalpy $\Delta H_e = D_i$ - D_f + $0.5\,(hL\nu_i$ - $hL\nu_f)$ including the energy difference of zero-point vibrations of broken and formed bonds (it

represents a change in the potential energy of the system). Here v_i is the frequency of vibration of the molecule along the broken bond, v_f is the frequency of vibration of the molecule along the formed bond, D_i is the bond dissociation energy of the broken bond, $D_{ei} = D_i + 0.5hLv_i$, D_f is the bond dissociation energy of the formed bond, $D_{ef} = D_f + 0{,}5hLv_f$.

2. The classical potential barrier of the activation E_e (1), which includes the zero-point energy of the broken bond.

3. The parameters $b = \pi(2\mu_i)^{1/2}v_i$ and $b_f = \pi(2\mu_f)^{1/2}v_f$, that describe the potential energy dependence of the atoms vibration amplitude along the breaking (*i*) and forming (*f*) valence linkage. $2b^2$ is the force constant of the linkage, μ_i is the reduced mass of atoms due to breaking bond, μ_f is the reduced mass of atoms due to forming bond.

4. The parameter r_e, which is the integrated stretching of breaking and forming bonds in the transition state.

5. The pre-exponential factor A_0 per equireactive bond in the molecule.

According to statistically determined value of br_e, based on formula (4), it is possible to estimate the value of classical potential barrier by the formula:

$$\sqrt{E_e} = \frac{br_e}{1-\alpha^2}\left(1-\alpha\sqrt{1-\left(1-\alpha^2\right)\frac{\Delta H_e}{(br_e)^2}}\right) \tag{6}$$

Thus, we can assume that the dependence of classical potential barrier E_e of thermochemical characteristics of reagents and kinetic characteristics of radical reactions can be represented as the functional relation:

$$E_e = F(\Delta H_e, T, nA_0, \alpha, br_e) \tag{7}$$

Then the task of ANN work for predicting the values of the classical potential barrier E_e as a functional relation of thermochemical and kinetic characteristics of reagents and reactionary center of radical reaction with subsequent calculation of activation energies and specific reaction rate by the formulas (1) and (2) reduce to approximation of unknown functional relation (7).

Problem Solution
Approximation of the classical potential barrier of radical bimolecular reaction in an artificial neural network

The experimental sample includes 113 radical reactions of alkyl, allyl and aryl radicals with various hydrocarbons, at that 37 radical reactions are control sample. Rate constants were obtained from a database information system [12], dissociation energy of C-H bonds [16].

In this paper the analysis of the experimental data suggests the presence of weak parabolic trend with pair correlation coefficient (R = 0.7874) and dispersion relation (F = 2.6071) Fisher, according to the enthalpy of radical bimolecular reactions from the square root of the value of the classical potential barrier (Fig.1):

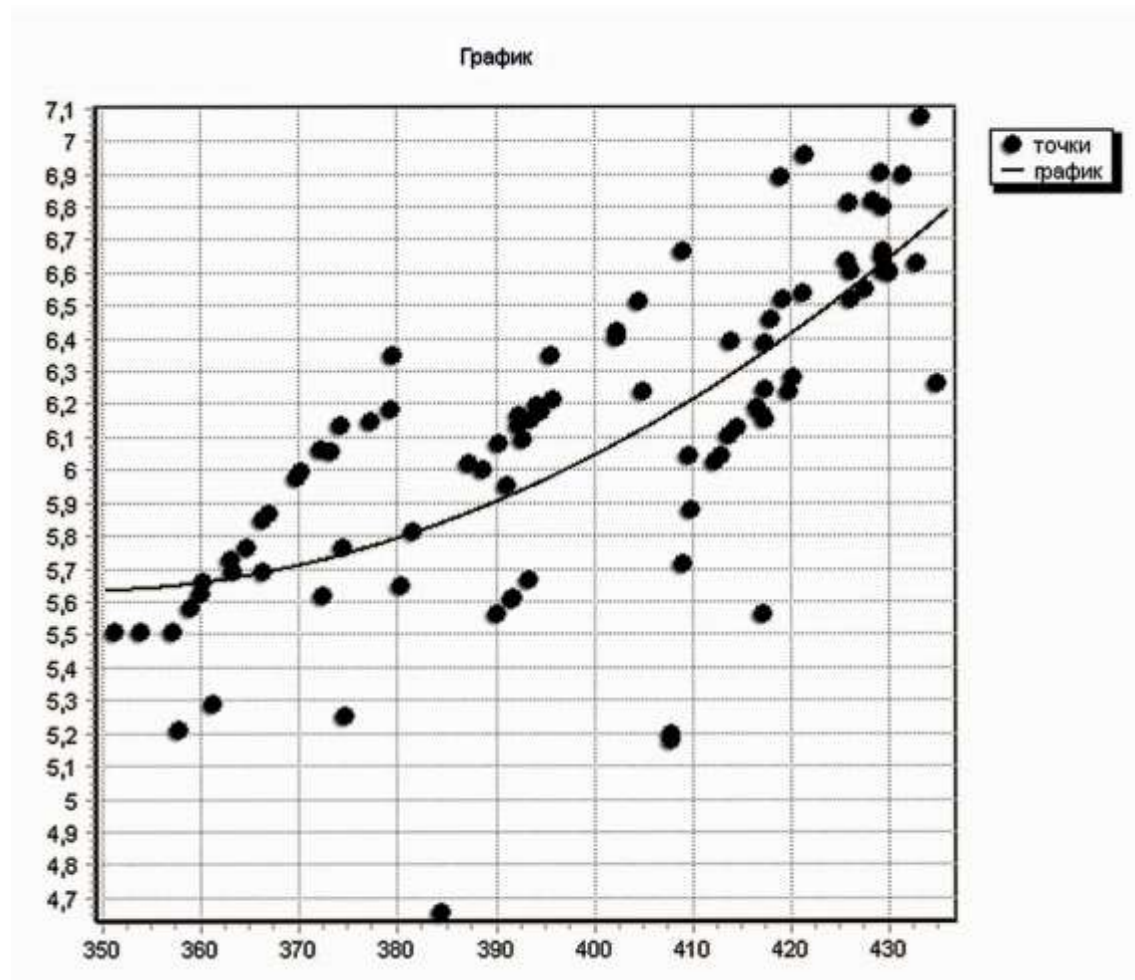$$\Delta H_e = 1.6 \times 10^{-4} E_e + 0.039\sqrt{E_e} + 8.172 \qquad (8)$$

Fig. 1 — Parabolic trend according to enthalpy of considered radical bimolecular reactions from the square root of the value of their classical potential barrier

This fact allows to use the kinetic parameter br_e as an experimental index of the reaction center of radical bimolecular reaction and to use this parameter in the ANN learning process.

So for reactions R + R_1H α=1, we will consider dependence

$$E_e = F(D_i, D_f, br_e) \qquad (9)$$

where all rate constants are counted for temperature 298 K. To approximate the dependence (9) we used feed-forward artificial neural network [17] with a typical architecture shown in Fig. 2. We used the ANN having 3 inputs, 1 inner layer (7 neurons) and 1 output.

ANN work is set by the formulae:

$$NET_{jl} = \sum_i w_{ijl} x_{ijl},$$

$$OUT_{il} = \Phi(NET_{jl} - \theta_{jl}),$$

$$x_{ij(l+1)} = OUT_{il},$$

$$\delta = 0.5 \sum_j \sum_k \left(y_j^k - d_j^k\right)^2.$$

(10)

where the index i will always denote the input number, j — the number of neurons in the layer, l — the number of the layer; x_{ijl} — the i-th input of the j-th neuron in the layer l; w_{ijl} — the weighting factor of the i-th input of the neuron number j in the layer l; NET_{jl} — the signal NET of the j-th neuron in the layer l; OUT_{jl} — the output signal of the neuron; θ_{jl} — the threshold of the neuron j in the layer l; x_{jl} — the input column vector of the layer l.
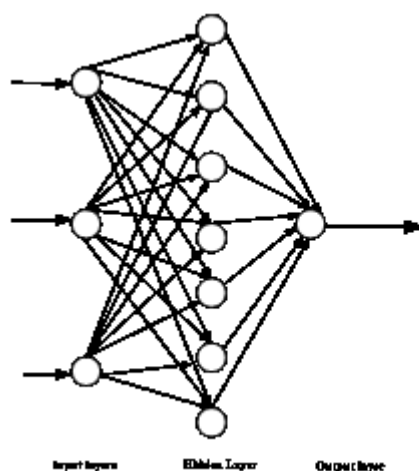

Fig. 2 — Typical architecture of feed-forward artificial neural network

The ANN input vector is set as the vector $x_0 = \{D_i, D_f, br_e\}$, output data is equal to E_e.

The method of back propagation of the error [8] was used as training procedure. Activation function is a sigmoid function and is set by the following formula:

$$f(x) = \frac{1}{1 + e^{-x}}$$

(11)

For the ANN training 11300 iterations were required on the training set of 113 samples, 37 samples were used for testing. Training set was constructed from the elemental radical reactions R + R₁H in the liquid

phase, where R - a alkyl, allyl or aryl radical and R_1H – a hydrocarbon molecule.

Then the analysis of the kinetic parameter values br_e was performed for various reactionary centers and the experimental index of reaction center $br_{e\text{-}class}$ was calculated, as shown in Table 1. We used $br_{e\text{-}class}$ calculated for the certain reactionary centers for prediction of the classical potential barrier.

Table 1 — An empirical index of the reactionary center and parameter A_0.

br_e_class kJ/mol	Reactionary centres (Reactions)		A_0
	Molecule	Radical	
	$-CH_2CH_3$	$-CH_2C°H_2$	10^9
14.93	$C°Cl_3$ + alkanes		
16.61	$C°H_3$ + cycloalkanes		
17.41	$C°H_3$ + CH_3CH_2CN		
16.16	$CH_3(CH_2)_7C°H_2$ +CH_3CN		
15.05	$C°Cl_3$ + cycloalkanes		
	$(-CH_2)_3CH$	$(-CH_2)_3C°$	10^9
15,01	$C°Cl_3$ + $(CH_3)_3CH$		
	$-CH_2CH=CHCH_3$	$-C°HCH=CHCH_3$	10^8
17.97	$C°H_3$ + alkenes		
18.61	$C°H_3$ + cycloalkenes		
17.52	$CH_3(CH_2)_4C°H_2$ +cycloalkenes		
16.37	*cyclo*-$[(CH_2)_2C°H]$ + cycloalkenes		
14.93	$C°F_3$ + alkenes		
17.13	$CF_3CF_2C°F_2$ + alkenes		
17.22	$C°Cl_3$ + cycloalkenes		
15.02	$C°Cl_3$ + $C_6H_5CH_2CH=CH_2$		
	$ArCH_3$	$ArC°H_2$	10^8
16.66	$C°H_3$ + arenes		
16.70	$(CH_3)_3C°$ + arenes		
18.47	$C_6H_4(C°H)C_6H_4$ + $C_6H_5CH_2C_6H_5$		
	$=CH(OH)$	$=C°(OH)$	10^9
16.35	$C°H_3$ + alcohols		
16.10	$C°F_3$ + alcohols		
15.35	$C°Cl_3$ + alcohols		
	$-CH(O)$	$-C°(O)$	10^9
15.61	$C°H_3$ + aldehydes		
15.41	$C°Cl_3$ + aldehydes		
	$-C(O)CH_2-$	$-C(O)C°H-$	10^8
16.44	$C°H_3$ + ketones		
15.41	$C°Cl_3$ + ketones		
	$-CH_2OCH_2-$	$-CH_2OC°H-$	10^9
15.98	$C°Cl_3$ + ethers		
17.06	$(CH_3)_3C°$ + cycloethers		
16.53	$-CH_2(O)OCH_2-$	$-C°H(O)OCH_2-$	10^9

| 16.22 | $-CH_2C(O)OH$ | $-C°HC(O)OH$ | 10^8 |

This effect, the ANN was retrained taking into account the obtained values of the experimental index of reactionary center. Learning outcomes for both ANN control sample are shown in Table. 2.

Table 2. shows the comparison of predictions of values of the classical potential barrier of the reaction using ANN (E_{ANN}), когда br_e was calculated on the formula (3), predictions of values of the classical potential barrier of the reaction using the ANN with an empirical index of the reactionary center (E_{ANN2}), when br_e was taken from Table 1, and the values of the classical potential barrier (E_e) calculated by the formula (1).

Table 2 - Training results of ANN

Reaction	E_e kJ/mol	E_{ANN} kJ/mol	E_{ANN2} kJ/mol
$C°Cl_3 + (CH_3)_3CH$	60.2	58.0	60.6
$CH_3(CH_2)_7C°H_2 + CH_3CN$	55.6	55.9	58.8
$C°Cl_3 + cyclo-[(CH_2)_6]$	62.2	59.3	64.1
$C°Cl_3 + cyclo-[(CH_2)_5]$	65.1	62.2	64.2
$C°H_3 + cyclo-[CH(CH_3)(CH_2)_5]$	47.4	48.5	51.7
$C°Cl_3 + CH_3(CH_2)_4CH_3$	66.5	63.1	66.5
$C°H_3 + CH_3CH_2CN$	51.7	52.5	55.1
$C°H_3 + cyclo-[CH(CH_3)(CH_2)_4]$	46.1	47.3	50.3
$C°Cl_3 + C_6H_5CH_2CH=CH_2$	38.4	37.7	40.6
$C°H_3 + CH_2=C(CH_3)_2$	44.8	45.1	47.5
$cyclo-[(CH_2)_2C°H] + cyclo-[CH=CH(CH_2)_4]$	30.4	29.6	31.4
$C°H_3 + cyclo-[CH=CH(CH_2)_6]$	46.5	46.4	49.0
$C°H_3 + cyclo-[CH=CH(CH_2)_5]$	45.3	44.5	44.5
$C°H_3 + cyclo-[CH=CH(CH_2)_3]$	44.5	43.4	42.7
$C°F_3 + CH_2=CHCH_2CH_3$	17.0	18.1	18.6
$C_6H_4(C°H)C_6H_4 + C_6H_5CH_2C_6H_5$	85.3	85.8	88.3
$CH_3(CH_2)_2C°H_2 + C_6H_5CH_2OCH_2C_6H_5$	49.9	49.8	52.4
$CH_3(CH_2)_4C°H_2 + C_6H_5CH_3$	49.1	49.7	51.6
$(CH_3)_3C° + C_6H_5CH_3$	60.2	59.8	60.5
$C°H_3 + C_6H_5CH_2CH_3$	36.5	36.6	39.4
$C°H_3 + C_6H_5CH(CH_3)_2$	33.8	33.1	33.8
$C°F_3 + (CH_3)_2CHOH$	38.8	38.9	43.0
$CF_3C°HCl + CH_3CH_2OH$	57.2	57.5	61.1
$C°HBr_2 + CH_3OH$	60.3	60.0	61.3
$C°Cl_3 + (CH_3)_2CHOH$	56.9	55.3	58.5
$C°H_3 + CH_3OH$	51.7	52.6	53.0
$C°H_3 + CH_3CH_2OH$	48.3	49.3	52.1
$CH_3(CH_2)_4C°H_2 + CH_3C(O)CH_3$	56.2	56.3	58.9
$C°H_3 + cyclo-[C(O)(CH_2)_5]$	43.9	44.6	49.6
$C°Cl_3 + CH_3CH_2C(O)CH_2CH_3$	61.4	59.3	62.4

$C°H_3 + CH_3CH_2C(O)CH_2CH_3$	41.0	41.5	44.4
$C°H_3 + (CH_3)_2CHC(O)CH(CH_3)_2$	37.0	37.1	40.2
$C°H_3 + cyclo\text{-}[C(O)(CH_2)_4]$	43.9	44.6	49.6
$(CH_3)_3C° + cyclo\text{-}[O(CH_2)_4]$	68.6	67.3	70.7
$C°H_3 + CH_3C(O)OCH_3$	37.3	37.2	39,9
$C°H_3 + CH_3CH_2C(O)OCH_3$	50.5	51.5	54.3
$C°H_3 + CH_3C(O)OH$	56.4	56.9	60.3

As seen in Table. 2, the ANN predicts with good accuracy the value of the classical barrier E_{ANN2} for reaction of alkyl, allyl or aryl radicals with hydrocarbons in the liquid phase. The maximum absolute error is 1.93 kJ / mol, and the minimum error is 1.25 kJ / mol, the mean square error for entire sample is 0.46 ± 0.42 kJ / mol, and these data are in agreement with error experimental methods of determining the activation energy of these reactions ± 4 kJ / mol.

Prediction of rate constants of reactions R + R₁H

Estimation of reaction rate constants R + R₁H in the liquid phase is based on the ormulas (1), (2) where the classical potential barrier is calculated using the ANN based on the dissociation energy of the broken bond and the experimental index of reactionary center. The general schematic of algorithm is shown in the Fig. 3. The algorithm uses the database that contains the pre-exponential factor for one equivalent reaction bond and the experimental index of reactionary center for various groups of hydrocarbons.
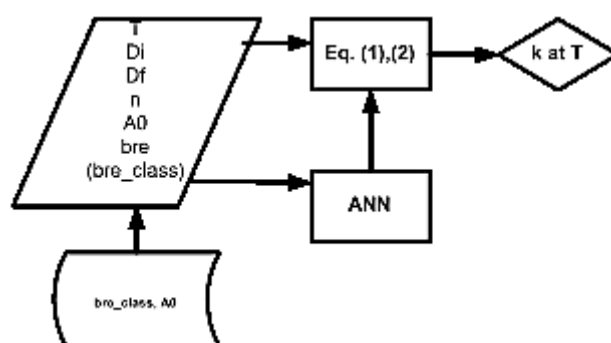

Fig. 3 — Basic scheme of hybrid algorithm of application of artificial neuronets for evaluation of rate constants of radical bimolecular reactions

In Figure 4 for the experimental rate constants k_{exp} and the calculated rate constants k_{cal} there is the diagram of $lg(k_{exp})$ on $lg(k_{cal})$, which is described by the linear correlation equation:

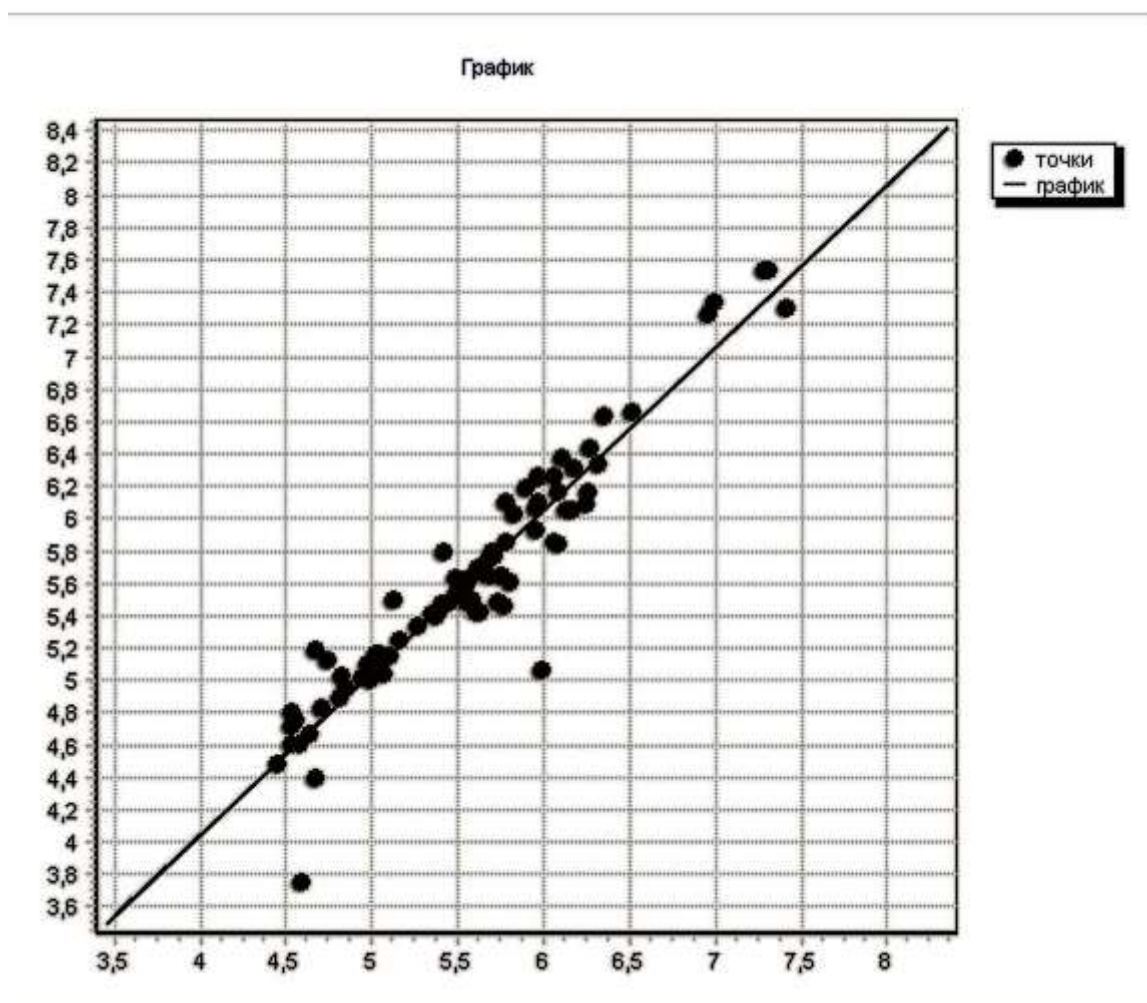$$lg(k_{exp}) = 1.01 + 3.50 \times 10^{-2} lg(k_{cal}) \tag{12}$$

Fig. 4 — Linear correlation dependence of lg(k$_{exp}$) on lg(k$_{cal}$).

The pair correlation coefficient for correlation ratio (12) is R = 0.998, and it indicates a good agreement between calculated and experimental values of logarithms of the rate constants. The difference between logarithms of the calculated and the experimental rate constants is anywhere from 1.09×10^{-3} to 0.34, and mean square error for the entire sample is 0.08 ± 0.07. The mean square error of relative error in calculation of rate constants is 18.5 %, and the maximum error is less than 35 %.

It is necessary to pay attention to the problem of computational accuracy in estimating the reactivity of organic molecules from the experimental data. The relative error in determination of the rate constants of radical bimolecular reactions is anywhere from 10 to 35 %, and it affects directly the quality of the experimental sample. The absolute error in determination of activation energy is from 2 to 4 kJ / mol. So, if the error is 3 kJ / mol in activation energy, this leads to the

relative error in calculation of the rate constant of 120 %, that is more than an order of magnitude.

Thus, the relative error in the estimation of rate constant through the activation energy of radical bimolecular reaction, calculated using the ANN (is trained on experimental data), averages 35%.

Conclusion

The hybrid algorithm of calculation of rate constants of bimolecular radical reactions on the experimental thermochemical data and an empirical index of the reactionary center with the use of an artificial neural network is offered.

The results of the prediction of reactivity of bimolecular radical reactions of hydrocarbons with hydrocarbon radicals in liquid phase are within the limits of the experimental error. The error of prediction of the classical potential barrier of the radical reaction using the ANN in control samples (of 39 samples) was 2.5 ± 1.5 kJ / mol, which is within the experimental error (± 4 kJ / mol).

The rate constants of $R + R_1H$ reactions in a liquid phase are estimated by proposed algorithm with relative error of approximately 35%, and it can serve as the good estimate of the reactivity of reagents in such reactions.

Thus, this developed method for estimating the reactivity of reactions of alkyl, allyl, aryl radicals with various hydrocarbons in the liquid phase using the ANN is a good theoretical implement in researching of such reactions.

References:
1. Gasteiger, J., Zupan, J. Neural networks in chemistry // Angev. Chem. Int. Ed. Engl. 1993. Vol. 32. — P. 503-527.
2. Himmelblau, D.M. Applications of Artificial Neural Networks in Chemical Engineering // Korean Z Chem. Eng. 2000. Vol. 17. N. 4. — P. 373-392.
3. El-Bakry, H.M., Mastorakis, N. A. new fast neural network model // Recent Researches in Applied Computers and Computational Science, ACACOS'12 Proceedings of the 11th WSEAS international conference on Applied Computer and Applied Computational Science. Rovaniemi, Finland. Appril 18-20, 2012. — P. 224-231.
4. Reinaldo, F., Siqueira, M. Camacho, R. Reis, L.P. Multi-Strategy Learning made easy // Proceedings of the 10th WSEAS International Conference on COMPUTERS. Vouliagmeni, Athens, Greece, July 13-15, 2006. — P. 278-284.

5. Slavova, A. Reaction-Diffusion Cellular Neural Network Models // Advanced Topics on Neural Networks, Proceedings9th WSEAS International Conference on NEURAL NETWORKS (NN'08), Sofia, Bulgaria, May 2-4, 2008. — P. 63-66.

6. Xu, Y., Yu, X., Zhang, S. QSAR Models of Reaction Rate Constants of Alkenes with Ozone and Hydroxyl Radical // J. Braz. Chem. Soc. 2013. V. 24. N. 11. — P. 1781-1788.

7. Dutot, A.-L., Rude, J., Aumont, B. Neural network method to estimate the aqueous rate constants for the OH reactions with organic compounds // Atmospheric Environment. 2003. V. 37. — P. 269–276.

8. Kovacs, B., Toth, J. Estimating Reaction Rate Constants with Neural Networks // World Academy of Science, Engineering and Technology. 2007. V. 26. — P. 13-17.

9. Ahmadi, M., Nekoomanesh, M., Arabi, H. New Approach in Modeling of Metallocene-Catalyzed Olefin Polymerization Using Artificial Neural Networks // Macromol. Theory Simul. 2009. V. 18. — P. 195–200.

10. Bas, D., Dudak, F.C., Boyacı, I.H. Modeling and optimization III: Reaction rate estimation using artificial neural network (ANN) without a kinetic model // Journal of Food Engineering. 2007. V. 79. — P. 622–628.

11. Tumanov, V., Gaifullin, G. Application of the Artificial Neural Networks for the Prediction of Reactivity of Molecules in Radical Reactions // Mathematical Modelling and Simulation in Applied Sciences. Proc. of the 3rd International Conference on Energy, Environment, Devices, Systems, Communications, Computers (INEEE '12). Rovaniemi, Finland. Appril 18-20, 2012. — P. 62-65.

12. 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. — P. 121-126.

13. Semenov, N.N. Chemical Kinetics and Chain Reactions. London, Oxford Uiv.press. — 1935. — 632 C.

14. Denisov, E.T. New empirical models of free radical abstraction reactions // Uspekhi Khimii. Vol. 66. 1997. — P. 953-971.

15. Denisov, E.T, Tumanov, V.E. Transition-State Model as the Result of 2 Morse Terms Crossing Applied to Atomic-Hydrogen Reactions // Zhurnal Fizicheskoi Khimii. Vol. 68. 1994. — P. 719-725.

16. Luo, Y.-R. Comprehensive Handbook of Chemical Bond Energies // CRC Press, Boca Raton, London - New York, 2007. — 1655 p.

17. Larose D.T. Data mining methods and models // John Wiley & Sons, Inc Publication. 2006. – 330 p.