Bonding parameter topological index application in lanthanides quantitative structure-activity relationship

Hanbiao Yang\(^1\), Lin Yang\(^2\) *, and Yingying Jin\(^2\)

\(^1\) School of Mathematics and Computational Science, Wuyi University, Jiangmen 520025, China.
\(^2\) Department of Chemical Materials, Jiangmen Vocational and Technical College, Jiangmen 520025, China.

*Corresponding author Email: lin61629694@163.com

Abstract. Quantitative structure relationship (QSAR) is an important way to explore chemical properties. This paper mainly studies the element bonding parameter topological index of lanthanides and the linear regression equation of its hydration properties. Based on the predecessors, this paper improves the algorithm of element bonding parameter topological index. A linear regression equation with a higher correlation coefficient was obtained. It provides a basis for further study of the theoretical characteristics of lanthanides hydration properties.

Keywords: Element bonding parameter topological index; Hydration properties; Lanthanides.

1. Research background

Quantitative structure relationship (QSAR) is an important research direction in chemistry. We mainly studied the relationship between the molecular structure and chemical activity of compounds. Quantitative structure-activity relationship (QSAR) has developed from individual and qualitative description to general and quantitative mathematical model expression. In this paper, we mainly discuss the use of the mathematical tool of element bond parameter topological index to describe the chemical molecular structure of compounds, and use computer software (matlab) to fit the linear relationship between the bonding parameter topological index and the hydration properties of lanthanides, that is, regression equation.

The hydration properties of lanthanides are an extremely important physiological chemistry. He mainly has lanthanides metal ionic hydration energy (H), hydrolysis constant ($PK_1$), hydroxide ($PK_m$) and Complex Stability Constant ($lgK$) parameters. He can be used to compare the strength of the hydration properties of lanthanides ionic.

In the Electrolyte aqueous solution, there is an electronically between the ionic and polar water molecule with an electric charge. There is a layer of water around the charged ionic. This layer of water attracts nearby water molecules under the action of the hydrogen bond. We call this phenomenon of ionic and water molecules combined as the hydration of ionic. The energy released in the hydration can be called hydration energy. Ionic hydration produces two effects, namely reducing the number of free water molecules in the solution, increasing the ionic volume, changing the ionic activity and...
conductivity; destroying the tetrahedral structure of the water, and also changing the dielectric constant of the ionic adjacent water molecule layer, weakened the mutual attraction of ionic. These effects can explain the anomalies of ionic conductance and ionic radius. Using this interpretation we can establish the theoretical basis for the topological index of the key parameters.

The hydrolytic equilibrium is a chemical equilibrium in which the equilibrium constant is the hydrolysis constant. For example, the hydrolysis constant of NaAc (Strong alkaline weak acid salt) is:

\[ K_h = \frac{K_w}{K_a} \]  

Where \( K_h \) is the hydrolysis constant, \( K_w \) is the ion product of water, and \( K_a \) is the weak acid ionization constant. It reflects the hydrolysis properties of the metal ion compound.

Complexation refers to the chemical reaction of metal ionic and complexing agents to generate metal complexes that are difficult to ionize and have high stability. The complexing agent is an organic compound containing salinization radicals (-COOH, -OH, =NOH or -SO3H), which can be replaced by a metal ionic under certain conditions; and a complexed atomic group (-N=N-, The O, S or N with a free electron pair in >C=O, >C=S, -NH\(_2\) or -NOH) is combined with the metal ionic with a coordinate bond. The resulting ring structure metal complex is called Complex salt. The complex salt has a low solubility, is insoluble in water, and has a vivid color. Complexation has high sensitivity and special effects, and is used for the determination of metal ionic in hygienic analytical chemistry. It has strong practicability. The complexing agent required for practice requires that the resulting complex must be stable and complex. If a variety of complexes are formed, the stability constant \( K \) of each complex will vary greatly. Complex Stability Constant is defined by \( K \) instability and has a reciprocal relationship with the instance constant \( K \). Complex Stability Constant can be used to indicate the stability of metal ionic and complexing agents to produce metal complexes that are difficult to ionize and have high stability. Has a strong practical application value.

Quantitative structure activity relationship (QSAR) research is an important research direction in chemometrics and is widely used in drug design methods. The quantitative structure activity relationship aims to establish a quantitative relationship between the physiological activity of a series of compounds and their physicochemical property parameters or structural parameters (including two-dimensional molecular structure parameters, three-dimensional molecular structure parameters, Illustrative structural diagrams, etc.) by reasonable mathematical statistics. Then, through these quantitative relationships, the corresponding characteristics of the compound are guessed, and the designer is guided to structurally modify the physiologically active substance, thereby greatly shortening the development cycle of the high-performance compound and saving the research and development cost.

The topological index research application is a combination of modern computational chemistry, structural chemistry and quantum chemistry [7] [8]. In particular, its good physicochemical property reflecting the molecular structure has become an important part of the quantitative structure activity relationship (QSAR), and it has received more and more attention from researchers. The topological index can represent a molecular map with a single topological index, commonly referred to as a topological invariant, and the closer the molecular topology is, the closer the value of the topological index is. To date, there have been more than forty kinds of topological indexes, among which the well-known topological index molecular connectivity indices [1] proposed by chemists Kier and Hall according to topological theory. After the key parameter topological index proposed by Xin Houwen, Wu Qixun et al., a certain quantitative structure activity relationship (QSAR) was also obtained in the organic substance study [2]. On the basis of this, we improved the algorithm of the element bonding parameter topological index in the lanthanides hydrational quantitative structure activity relationship (QSAR), and further improved its linear dependence.
2. Element bonding parameter topological index
The general form of the key parameter topological index is

$$H_e = \left[ \sum e^{\frac{1}{(V_i - Y_i)^{\frac{1}{2}}}} \right]^2$$

(2)

In formula (1), $p_i$, $q_i$ is the degree of the apex atom at the $i$th side of the $i$thive structural diagrams, $V_i$, $Y_i$ is the key parameter corresponding to the two vertex atoms, $i$ is the sum of all the edges in Illustrative structural diagrams, $t$ represents the type of the vertex atomic key parameter.

We further simplify this formula and expand the $e^{(V_i - Y_i)}$ term in equation (2), taking only the first two items:

$$H_e = \left[ \sum \frac{1}{[1 + (V_i - Y_i)]^{\frac{1}{2}}p_i q_i} \right]^2$$

(3)

We continue to further simplify (3). For an atom of an element, it corresponds to the $i$-edge of the molecular graph degenerating into a point, so it can be replaced by the atomic parameter R ($V_i - Y_i$), the atomic parameter S is substituted for $p_i$, $q_i$, and the following is defined by HE as the element bonding parameter Topological index:

$$H_E = \frac{1}{(1 + R)^{\frac{1}{2}}S}$$

(4)

The reciprocal which still contains a somewhat priced square root in formula (4) is consistent with the main meaning of the widely used topological index topological index molecular connectivity indices.

3. Topological invariant of the lanthanide’s key parameter topological index
In equation (4), the binary function $H_E$ is an increasing function with respect to the argument atomic parameters R and S. When R is taken as the metal radius of lanthanides, and S takes the Valence electron N of the lanthanide’s atom, R becomes strictly smaller as the lanthanides atomic number grows from small to large. However, S gradually increases as the atomic number of lanthanides grows from small to large. From this perspective, equation (4) is not a continuous topological invariant that satisfies the topological index. But the data of lanthanides is discrete, we can know (4) the first term of the denominator on the right side of the equal sign, ie $(1+R)$, in the lanthanides with the element number from small to large, $(1-R)$ the largest span of decline the decline is 97.24%. And (4) the first term of the denominator on the right side of the equal sign, that is, in the lanthanides as the element number grows from small to large, the rise span is minimally up by 104%. Even in the most extreme case, the overall (4) formula rose by 100.16%. That is to say, under the discrete data of lanthanides, with the lanthanides atomic number from small to large [9], the inference (4) type topological index is a strict reduction function, that is, under lanthanides discrete data, (4) has a topological invariant. Of course, this property can also be realized in the actual calculated discrete data. The element bonding parameter topological index of lanthanides is strictly from large to small as the element number is from small to large. This conclusion eliminates the interference of the two outlier elements of Eu and Yb.

4. Algorithm improvement in this paper
In order to study the physical chemistry of an element, we must specifically define the parameters of R, S in equation (4). In the literature [4], the researchers took R as the lanthanides Ln$^{3+}$ ionic radius, and S took the Valence electron N of the lanthanide’s atom. For example, the Ln$^{3+}$ ionic radius for lanthanides La is 1.06Å, the Valence electron is 3, and the substitution (4) is calculated to give $H_{E1}$=0.28.

However, this paper has improved the algorithm for element bonding parameter topological index. Take R as the lanthanides radius. For example, for Lanthanides La, the Ln$^{3+}$ ionic radius is 187.7, the Valence electron is 3, and the equivalent substitution (4) is calculated as $H_{E2}$=798.
In particular, to change \( R \) from \( \text{Ln}^{3+} \) ionic radius to lanthanides metal radius, you need to do an operation to eliminate outlier. Most of the Ionics of lanthanides show a tendency to decrease uniformly with the atomic number, but in Eu, the two elements of \( \text{Yb} \) largely violate this rule. If you do not eliminate the interference of these two elements, it will lead to the correlation coefficient will be lower than 0.9 in the calculation. Therefore, in the calculation process of this paper, the metal radius data of the two elements \( \text{Eu} \) and \( \text{Yb} \) are first removed. Then, the metal radius data of the \( \text{Eu} \) and \( \text{Yb} \) elements are "completed" by the fill missing function in MATLAB.

After obtaining the new element bonding parameter topological index for each lanthanide, we can get the linear regression equation of its hydration properties.

\[
P = a + bH_E
\]  

(5)

Where \( P \) is a certain hydration property of lanthanides, \( a, b \) is a parameter of the linear regression equation. The coefficient of the linear regression equation for each hydration property of lanthanides is shown in Table 1.

Tab. 1 The linear regression equation’s coefficient for each hydration property of lanthanides

| H/Kcal\(^\text{mod-1}\) | \(P_K1\) calculated value | \(P_Km\) calculated value | \(\lg K\) calculated value |
|-------------------------|--------------------------|--------------------------|--------------------------|
| a                       | 959.26                   | 7                        | 7.13                     |
| b                       | -56646                   | 673.62                   | -708.47                  |

Through the above data processing techniques, we can get a linear correlation equation with higher correlation coefficient than the literature [4] [5]. A higher linear correlation coefficient can indicate that the regression equation can reflect the relevant chemical properties better and more linearly. For example, in [4], [5], for the linear regression equation of the H/Kcal\(^\text{mod-1}\) lanthanide example hydration energy property, the correlation coefficient is 0.977, and in this calculation, the hydration energy property of the H/Kcal\(^\text{mod-1}\)lanthanides example is in the linear regression equation, the correlation coefficient is 0.987. That is to say, using the lanthanides metal radius as an atomic parameter, the regression equation can be used more accurately to predict the physical chemistry index of lanthanides.

Specifically, the hydrolysis constant \(P_{K1}\) of lanthanides ionic in the original algorithm, the correlation coefficient is 0.965, this paper is 0.977, slightly improved, the correlation of the hydroxide \(P_{Km}\) in the original algorithm, the correlation coefficient is 0.946, this paper is 0.946, exactly the same. In the original algorithm, the EDTA stability constant \(\lg K\) is 0.937, which is 0.957 in this algorithm. Therefore, it can be seen that in the lanthanide’s hydration properties such as H/Kcal\(^\text{mod-1}\) and \(\lg K\), the algorithm can improve the correlation coefficient and obtain a stronger linear dependence. It indicates that H/Kcal\(^\text{mod-1}\) and \(\lg K\) have a higher linear correlation with the metal radius of lanthanides than the ionic radius. The hydration properties of each lanthanides are shown in Table 2.

Tab. 2 hydration parameters of each lanthanides

| Lanthanides and electronic numbers | Metal radius | Lanthanides key parameter | \(H/\text{Kcal}^\text{mod-1}\) Calculated | \(H/\text{Kcal}^\text{mod-1}\) Literature value | \(P_{K1}\) Calculated | \(P_{K1}\) Literature value | \(P_{Km}\) Calculated | \(P_{Km}\) Literature value | \(\lg K\) Calculated | \(\lg K\) Literature value |
|----------------------------------|--------------|--------------------------|------------------------------------------|------------------------------------------|----------------------|--------------------------|----------------------|--------------------------|----------------------|--------------------------|
| La 3                             | 187.7        | 0.00305                  | 786.7                                   | 798                                      | 9.1                   | 9                        | 5.9                   | 5.1                      | 15.2                  | 15.5                     |
| Ce 4                             | 182.5        | 0.00272                  | 805.7                                   | 813                                      | 8.8                   | 8.8                      | 5.2                   | 5.3                      | 15.9                  | 15.9                     |
| Pr 5                             | 182.8        | 0.00243                  | 822.2                                   | 821                                      | 8.6                   | 8.6                      | 5.4                   | 5.3                      | 16.5                  | 16.4                     |
| Nd 6                             | 182.1        | 0.00222                  | 833.7                                   | 826                                      | 8.5                   | 8.4                      | 5.6                   | 5.5                      | 16.9                  | 16.6                     |
| Pm 7                             | 181.0        | 0.00207                  | 842.4                                   | 8.4                                      | 5.7                   | 5.6                      | 7.3                   | 7.3                      | 17.3                  | 17.4                     |
| Sm 8                             | 180.2        | 0.00195                  | 849.5                                   | 8.3                                      | 5.8                   | 8.3                      | 5.7                   | 17.5                     | 17.4                  | 17.4                     |
| Eu 9                             | 204.2        | 0.00172                  | 960.0                                   | 7.0                                      | 8.3                   | 7.1                      | 5.8                   | 21.7                     | 17.3                  | 17.3                     |
| Gd 10                            | 180.2        | 0.00172                  | 861.1                                   | 844                                      | 8.2                   | 8.4                      | 5.9                   | 5.8                      | 18.0                  | 17.3                     |
| Tb 11                            | 178.2        | 0.00168                  | 864.7                                   | 8.1                                      | 8.2                   | 5.9                      | 8.1                   | 18.1                     | 17.9                  | 17.9                     |
| Dy 12                            | 177.5        | 0.00161                  | 868.3                                   | 873                                      | 8.1                   | 8.1                      | 6.0                   | 5.8                      | 18.2                  | 18.3                     |
| Ho 13                            | 176.6        | 0.00156                  | 871.5                                   | 873                                      | 8.1                   | 8.1                      | 6.0                   | 5.9                      | 18.4                  | 18.7                     |
| Er 14                            | 175.7        | 0.00151                  | 874.3                                   | 878                                      | 8.0                   | 8.0                      | 6.1                   | 6.1                      | 18.5                  | 18.8                     |
| Tm 15                            | 174.6        | 0.00147                  | 876.7                                   | 878                                      | 8.0                   | 8.0                      | 6.1                   | 6.1                      | 18.6                  | 19.3                     |
| Yb 16                            | 173.4        | 0.00139                  | 881.2                                   | 906                                      | 7.9                   | 7.9                      | 6.1                   | 6.1                      | 18.7                  | 18.7                     |
| Lu 17                            | 173.4        | 0.00139                  | 881.2                                   | 906                                      | 7.9                   | 7.9                      | 6.1                   | 6.1                      | 18.7                  | 18.7                     |
5. Summary
Using topological index for quantitative structure activity relationship research, we can measure the physicochemical property of the element by topological index, obtain its quantitative structure activity relationship with its chemical properties, and then guess the corresponding characteristics of the compound through these quantitative relationships. Therefore, the development cycle of high-performance compounds is greatly shortened, and research and development costs are saved [10] [11] [12]. According to the new algorithm in this paper, compared with the previous results, the regression equation of lanthanides hydration properties with higher coefficient can be obtained, which helps to provide the prediction accuracy of hydration properties research. It provides a certain pavement for the theoretical study of the properties of hydration [6]. And in the hydration energy H/Kcal mol\(^{-1}\) and EDTA stability constant \(\lg K\), the linear regression equation is more stable, more significant, and the data is more reliable, reflecting the hydrating properties of lanthanides. The method data in this paper is easy to obtain, and the physical meaning is clear. Although not all data, some of the results are improved compared with the old algorithm in linear correlation and other statistical indicators, which is worthy of application promotion.

Acknowledgments
This article is from The Doctor Project of Natural Science Foundation of Guangdong Province (No.2016A030310002), The Natural Science Foundation Youth Project of China (NO.11601393); National Natural Science Foundation of China: The quaternionic manifolds and quasiconformal mappings (No:11871379),The Innovation Project of Department of Education of Guangdong Province: The geometric and topological characteristics of quaternionic hyper-bolic manifold, The Innovation Project of Department of Education of Guangdong Province: Compilation of Linear Algebra Textbook Guided and Combining Engineering by OBE and The Innovation Project of Department of Education of Guangdong Province: Application of Topological Index and Other Mathematical Fields in Exploring Chemical and Physical Properties of New Energy Lanthanide Oxides.

References
[1] Kier L B, Hall L H. Molecular Connectivity in Chemistry and Drug Research. Academic Press, 1976, p12-19
[2] Zhang Hongguang, Xin Houwen. The key parameter topological index and its heterogeneous atoms in the related XY, XY2, XY3, XY4 isoforms. Chinese Journal of Chemical Physics, Vol. 6 (1989) No.2, p. 413-419.
[3] Wu Qixun, Qi Zhengxing, Pan Guoqing, et al. The topological index and application of lanthanides. Chemical Bulletin, Vol. 4 (1998) No.15, p. 103-105.
[4] Zhu Xian, Chen Liqin. Topological Index of Bond Parameters and Standard Electrode Potential of Common Elements. Chinese Journal of Applied Chemistry, Vol. 1 (2000) No.17, p. 32-35.
[5] Jin Jianhua. Correlation between topological index and lanthanidesphicochemical property. Journal of Wuhan University of Technology, Vol. 3 (2008) No.30, p. 15-19.
[6] XUE Ni, JIANG Jun-cheng, PAN Yong. Performance prediction of CO2 capture by imidazole ionic liquid based on novel topological index. Journal of Safety and Environment, Vol. 1 (2017) No.14, p. 78-81.
[7] Hanbiao Yang, Katsuro Sakai, Katsuhisa Koshino. A function space from a compact metrizable space to a dendrite with the hypo-graph topology. Open Mathematics, formerly Central European Journal of Mathematics, Vol. 1 (2015) No.13, p. 211-228.
[8] Hanbiao Yang. Zhongqiang Yang. Y, Zheng. Topological classification of function spaces with the Fell topology IV. Topology and its Applications. Vol. 2 (2017) No.28, p. 222-235.
[9] Hanbiao Yang, Liu Yang. A topological position of the set of strongly discontinuous maps in the set of upper semi-continuous maps. Quantitative Logic and Soft Computing, Ed. by Guojun Wang, Bin Zhao and Yongming Li, World Scientific Proceedings Series on Computer
[10] YANG Lin, LUO Linghong, WU Yifan, SHI Jijun, CHENG Liang, HU Xiao-qi. Preparation of GDC nanopowders by different methods and their application as barrier layers for SOFC cells. Journal of Ceramics, Vol. 2 (2014) No.35, p. 182-187.

[11] YANG Lin, DENG Hong, CHEN Yunjing. Research Status of Lead-free Piezoelectric Ceramics Based on Calcium Zirconate Titanate (BCZT). Journal of Ceramics, Vol. 3 (2016) No.37, p. 230-235.

[12] Yang Lin. Preparation and optimization of GDC powder and its application as a barrier layer for SOFC cell. Jingdezhen Ceramic Institute, 2014, p.15-25.