J. Basic. Appl. Sci. Res., 2(2)1619-1622, 2012

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Application of Novel Method for Investigation of Retention Behavior of Toxic Compounds in Wastewater

Hamid Reza Lotfizadeh[1], Hamzeh Karimi[2], Kamyar Arman[1], Hadi Nourizadeh[3]* and Abbas Farmani[3]

[1]Faculty of Science, Ilam Branch, Islamic Azad University, Ilam, Iran
[2]Faculty of Sciences, South Tehran Branch, Islamic Azad University, Tehran, Iran
[3]Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran

ABSTRACT

Wastewater is any water that has been adversely affected in quality by anthropogenic influence. It comprises liquid waste discharged by domestic residences, commercial properties, industry, and/or agriculture and can encompass a wide range of potential contaminants and concentrations. In the most common usage, it refers to the municipal wastewater that contains a broad spectrum of contaminants resulting from the mixing of wastewaters from different sources. A data set consisting of compounds in wastewater have been used to develop a quantitative structure retention relationship (QSRR) for their retention time. These data obtained by gas chromatography–mass spectrometry (GC-MS). QSRR model are usually obtained by splitting the data into two sets including training and test. Genetic algorithm and partial least square (GA-PLS) technique was used for QSRR model. The results indicate that GA-PLS can be used as an alternative modeling tool for QSRR.

KEY WORDS: *Diosorea zingiberensis* C.H. Wright (DZW); Diosgenin; GC-MS; Wastewater; Genetic algorithm-partial least square

INTRODUCTION

Wast Water or Wastwater is a lake located in Wasdale, a valley in the western part of the Lake District National Park, England. The lake is approximately 4.6 kilometres (almost 3 miles) long and 600 metres (more than a third of a mile) wide. It is the deepest lake in England at 79 metres (258 feet), and is owned by the National Trust. It is one of the finest examples of a glacially 'over-deepened' valley. The surface of the lake is about 200 feet above sea level, while its bottom is over 50 feet below sea level.

The head of the Wasdale Valley is surrounded by some of the highest mountains in England, including Scafell Pike, Great Gable and Lingmell. On September 9, 2007, Wast Water was announced as the winner of a vote to determine Britain's Favourite View by viewers of ITV. The steep slopes on the south eastern side of the lake, leading up to the summits of Whin Rigg and Illgill Head, are known as the Wastwater Screes or on some maps as The Screes. These screes formed as a result of ice and weathering erosion on the rocks of the Borrowdale Volcanic Group, that forms the fells to the east of the lake, towards Eskdale. They are approximately 2,000 feet, from top to base, the base being about 200 feet below the surface of the lake. A popular path runs the length of the lake, through the boulders and scree fall at the base of this craggy fell-side. On the north western side are the cliffs of Buckbarrow (a part of Seatallan) and the upturned-boat shape of Yewbarrow. Wast Water is the source of the River Irt which flows into the Irish Sea near Ravenglass.

Wastwater comes from Wasdale plus English water. The valley is pronounced as in was, not with a hard a: the name of the lake similarly. The lake is named Wast Water on Ordnance Survey maps but the spelling Wastwater is used with roughly equal frequency, including by its owner, the National Trust, along with the Cumbria Tourist Board, and the Lake District National Park Authority.

The Wasdale Lady in the Lake, Margaret Hogg, was murdered and her body was disposed of in the lake. She was found after eight years, with her body preserved like wax due to the lack of oxygen in the water. In February 2005 it was reported that a gnome garden complete with picket fence was removed from the bottom of Wastwater after three divers died in the late 1990s. It is thought the divers spent too much time too deep searching for the ornaments. Police divers report there is a rumour that the garden has returned at a depth beyond which they are allowed. PC Kenny McMahon, a member of the North West Police Underwater Search Unit, said Wastwater is quite clear at the bottom, but there's nothing to see. At a depth of about 48m, divers had taken gnomes down and put a picket fence around them. But several years ago there were a number of fatalities and the Lake District National Park Authority asked us to get rid of them. We went down there, put them in bags and removed the lot. But now there's a rumour about a new garden beyond the 50m depth limit. As police divers we can't legally dive any deeper so, if it exists, the new garden could have been purposefully put out of our reach.

***Corresponding Author:** Hadi Nourizadeh, Department of Chemistry, Faculty of Science, Arak Branch, Islamic Azad University, Arak, Iran, Tel.: + 989183452507 E-mail address: nourizadehadi@yahoo.com

Diosorea zingiberensis C.H. Wright (DZW) tuber is one of the widely-used raw materials in the pharmaceutical industry for diosgenin production, a chemical used for the synthesis of adrenal cortex hormone, sex hormone, progestational hormone and anabolic steroid. DZW tubers contain 1–5% saponin, 45–50% starch and 40–50% cellulose. Saponins mainly contain carbohydrate and diosgenin. The carbohydrates which connect with diosgenin by glucosidic bond are: d-glucose, d-xylose, d-galactose, l-rhamnose, l-arabinose, etc. In acidic condition, saponins could be hydrolyzed into diosgenin and monosaccharide.

Quantitative structure–retention relationship (QSRR) is statistically derived relationships between chromatographic parameters and descriptors related to the molecular structure of the analytes. A number of reports deals with QSRR retention calculation of essential oils compounds have been published in the literature [8-10]. There is a trend to develop QSRR from a variety of methods. In particular, genetic algorithm (GA) is frequently used as search algorithms for variable selection in chemometrics and QSRR. GA is a stochastic method to solve the optimization problems defined by fitness criteria, applying the evolution hypothesis of Darwin and different genetic functions, i.e. crossover and mutation [11, 12]. Partial least square (PLS) is the most commonly used multivariate calibration method [13, 14]. In the present study, GA-PLS was employed to generate QSRR model that correlate the structure of compounds in wastewater; with observed RT.

MATERIALS AND METHODS

Data set

Retention time of the 42 compounds in wastewater was taken from literature [15]. The data set was randomly divided into two groups including training set (calibration and prediction sets) and external (test) sets, which consists of 32 and 10 molecules, respectively.

Wastewater analysis

GC–MS analysis

GC–MS was employed to analyze organic compounds by liquid–liquid extraction using CH2Cl2 (chromatogram pure grade, Fisher Corporation, USA). Before extracting, the wastewater was adjusted to pH 6–7 and centrifuged (10,000 rpm, 15min) to remove the precipitates. Then 50ml supernatant was extracted in a separatory funnel using CH2Cl2 and the extraction was repeated three times. The CH2Cl2 phase was collected and dried using nitrogen. The 1ml pretreated sample was analyzed by 5973N/A2973 GC–MS system (Agilent Corporation, USA). High purity He (99.999%) was employed as carrier gas at the flow rate of 1ml/min. A DB-MS5 capillary columnwith inner diameter 0.25mmand length 30m was adopted in the separation system. The tempera- ture of the gasification compartment was first set to 40 ◦C for 5min, then raised linearly to 280 ◦C at the rate of 5◦C/min, and maintained at 280 ◦C thereafter. The electron energy was 70 eV and the electron double voltage was 1050V. Molecular weight was scanned in the range from 45 to 450Da. Organic compounds were analyzed with reference to the NIST02mass spectral library database. Those compounds which covered more than 0.5% of the integration area by GC–MS were taken as the primary organic constituents.

Computer hardware and software

All calculations were run on a HP Laptop computer with AMD Turion64X2 processor with windows XP operating system. The optimizations of molecular structures were done by the HyperChem 7.0 (AM1 method) and descriptors were calculated by Dragon Version 3.0 software's. Cross validation, GA-PLS and other calculation were performed in the MATLAB (Version 7, Mathworks, Inc.) environment.

Cross validation technique

Cross validation is a popular technique used to explore the reliability of statistical model. Based on this technique, a number of modified data sets are created by deleting in each case one or a small group (leave-some-out) of objects. For each data set, an input–output model is developed, based on the utilized modeling technique. Each model is evaluated, by measuring its accuracy in predicting the responses of the remaining data (the ones or group data that have not been utilized in the development of the model) [16]. In particular, the leave group out (LGO) procedure was utilized in this study.

Data pre-processing

Each set of the calculated descriptors was collected in a separate data matrix Di with a dimension of (m×n), where m and n are being the number of molecules and the number of descriptors, respectively. Grouping of descriptors was based on the classification achieved by Dragon software. In each group, the

calculated descriptors were searched for constant or near constant values for all molecules and those detected were removed. Before applying the analysis methods, and due to the quality of data, a previous treatment of the data is required. Scaling and centering is one of the pre-processing methods we need before performing the regression methods combined with FE. The results of projection methods depend on the normalization of the data. Descriptors with small absolute values have a small contribution to overall variances; this biases towards other descriptors with higher values. With appropriate scaling, equal weights are assigned to each descriptor, so that the important variables in the model can be focused. In order to give all variables the same importance, they are standardized to unit variance and zero mean (autoscaling).

RESULTS AND DISCUSSION

Results of the GA-PLS model

To reduce the original pool of descriptors to an appropriate size, the objective descriptor reduction was performed using various criteria. Reducing the pool of descriptors eliminates those descriptors which contribute either no information or whose information content is redundant with other descriptors present in the pool. The remained descriptors were employed to generate the model with the GA-PLS program. The best model is selected on the basis of the highest square correlation coefficient (R2) of prediction and simplicity of the model. The best GA-PLS model contains 21 selected descriptors in 13 latent variables space. The R2 for training and test sets was (0.784, 0.568), respectively. For this in general, the number of components (latent variables) is less than number of independent variables in PLS analysis. Plots of predicted RT versus experimental RT values by GA-PLS for training and test set are shown Fig. 1. Obviously, there is a close agreement between the experimental and predicted RT and the data represent a very low scattering around a straight line with respective slope and intercept close to one and zero. The PLS model uses higher number of descriptors that allow the model to extract better structural information from descriptors to result in a lower prediction error. Inspection of the results reveals a higher R2 value parameter for the training and test sets GA-PLS. The GA-PLS linear model has good statistical quality with low prediction error.

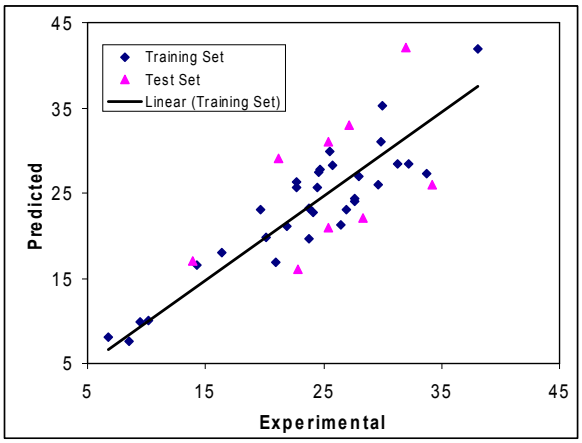

Fig.1. Plot of predicted RT obtained by GA-PLS against the experimental values

Conclusion

In this research, an accurate QSRR model for estimating the retention time of compounds in wastewater was developed by employing the GA-PLS technique. This model has good predictive capacity and excellent statistical parameters. It is easy to notice that there was a good prospect for the GA-PLS application in the QSRR modeling. It can also be used successfully to estimate the RT for new compounds or for other compounds whose experimental values are unknown.

REFERENCES

1. Marion Gust, Thierry Buronfosse, Olivier Geffard, Raphael Mons, Herve Queau, Jacques Mouthon, Jeanne Garric, Water Research, 44, 2010, 4517-4528

2. Vânia Figueira, Ivone Vaz-Moreira, Márcia Silva, Célia M. Manaia, Water Research, 45, 2011, 5599-5611

3. Guangyong Zhu, Xian Zhu, Qi Fan, Xueliang Wan, Resources, Conservation and Recycling, 55, 2011, 409-416

4. Niklas Janzen, Stefan Banzhaf, Traugott Scheytt, Kai Bester, Chemosphere, 77, 2009, 1358-1365

5. Koichi Terasaka, Ai Hirabayashi, Takanori Nishino, Satoko Fujioka, Daisuke Kobayashi, Chemical Engineering Science, 66, 2011, 3172-3179

6. Melanie Beck, Michael Radke, Chemosphere, 64, 2006, 1134-1140

7. Juan-qin XUE, Wei-bo MAO, Yu-jie WANG, Jing-xian LI, Ming WU, Xi LU, Transactions of Nonferrous Metals Society of China, 20,2010, 153-158

8. Chen, J.; Yang, T.; Cramer, S. M. J. Chromatogr. A. 2008, 1177, 207.

9. Noorizadeh, H.; Farmany, A. Chromatographia. 2010, 72, 563.

10. H. Noorizadeh, S. Sobhan Ardakani, T. Ahmadi, S. S. Mortazavic, M. Noorizadehd, Drug Test Anal, 2011, in press

11. Hemmateenejad, B.; Javadnia, K.; Elyasi, M. Anal. Chim. Act, 2007, 592, 72.

12. J. Aires-de-Sousa, M.C. Hemmer, J. Casteiger, Prediction of H-1 NMR chemical shifts using neural networks, Anal. Chem. 74 (2002) 80–90.

13. Hadi Noorizadeh, Mehrab Noorizadeh , Adnan S. Mumtaz, Journal of Saudi Chemical Society, 2011, in press

14. H. Noorizadeh, A. Farmany, Drug Test Anal, 2011, in press.

15. Linli Qiua, Hai Niub, Wen Huang, chemical engineering research and design 89 (2011) 239–247

16. H. Noorizadeh, A. Farmany, M. Noorizadeh, Quim. Nova, 34 (2011) 242-249.