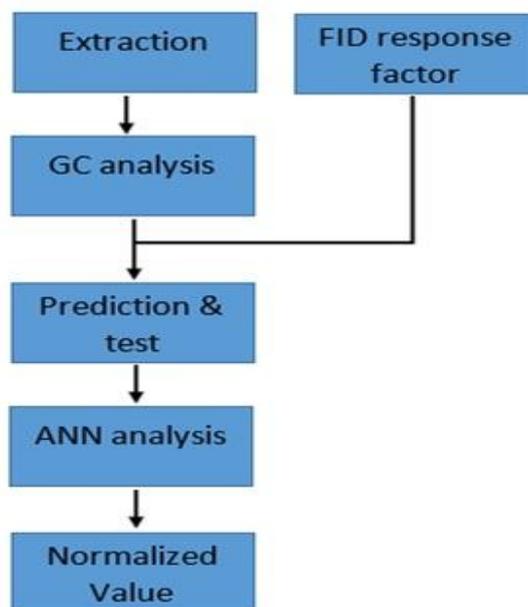


The Prediction of Adsorption Properties of a Solid Phase-Microextraction Fiber by Chemometrics methods

Mostafa Jafari*

ARTICLE INFO	ABSTRACT
Received: 29 January 2019 Revised: 4 February 2019 Accepted: 5 February 2019 Available online: 4 March 2019 DOI: 10.29088/SAMI/AJCA.2019.2.2032015	A new method for estimation of adsorption properties of a solid phase micro extraction fiber by artificial neural network (ANN) has been studied for the first time ever. An etched steel fiber which is simple prepared and durable was selected and adsorption of 12 analytes that were in four different chemical categories, was studied. 9 of them were selected as the training and 3 as the test. The amount of adsorptions were obtained through the direct extraction from aqueous and then GC analysis. The adsorption were analyzed by ANN. The results are quite satisfactory and the mean absolute percentage error of tests was 18.0 %. The method was simple, practical, straightforward, economical, and accurate. The method did not require many analyses.
KEYWORDS Artificial neural network Solid phase Micro extraction Adsorption Prediction	

GRAPHICAL ABSTRACT



* Corresponding author's E-mail address: ms.jafari@sbu.ac.ir, Tel.: +98 21 29903091
Faculty of Chemistry and Petroleum Sciences, Shahid Beheshti University, G.C, 1983969411, Tehran,
Islamic Republic of Iran.

Introduction

Solid Phase Micro Extraction (SPME) is a new sample pretreatment technique, which is green and solvent-less, simple and rapid [1]. In this method, a thin layer of absorber – mounted on a supporter- is used to extract the analytes. SPME is used in many fields, such as determination of environmental [2, 3] and biological samples [4, 5]. Many other features of SPME have been reviewed in somewhere else [5, 6].

While the theory for SPME has been exhaustively developed and many fibers presented, prediction of the adsorption capability of a fiber is often experimentally. The prediction of the capability of these coating to adsorb an especial species or a group of analytes is not truly clear. Experiment and some rule of thumb are just available methods to guess the adsorption character of a fiber.

Artificial neural network (ANN), a non-linear classifier, can be considered as one of the most important developing tools in multivariate analysis [7]. It is a parallel distributed processor that are able to extract information from a complicated data set, and then making ready to use. Superior classification and prediction capabilities of ANN make it a very strong tool to make an empirical model from data set. Recently, many works have been done with ANN to

find the relation between physico-chemical properties of experiment/material and instrumental response or final properties [8, 9].

The multilayered perceptrons (MLPs) are the most commonly used types of artificial neural networks. The choice of proper algorithm is the first challenges. Many algorithms are presented for MLP, some of them include bath back propagation (BBP), incremental back propagation (IBP), genetic algorithm (GA), quick prop (QP) and levenberg-marquardt (LM) are considered at 3.3.1. Over fitting is another problem, which is happen when data set is very small. On the other hand, the larger networks may result more complex functions. However, no proper technique exists to find how large a network should be for a specific application, prior the experiment.

Despite of uncountable papers in field of chemometrics, many chemometrics methods are not popular for many chemists. Many chemists do not know applications or cannot use chemometrics techniques in their fields, so they do not use chemometrics. On the other words, there is a gap between theories and routine applications of chemometrics for two reasons 1-difficulty of methods 2-lack of knowledge. Development of available methods for daily usages of all chemists can

fill this gap.

The goal of this study is to explain how chemometrics can be used to predict adsorption of a compound on a surface specifically a SPME fiber. Some different compounds from four categories were selected and extracted with a fiber. The extraction condition was optimized for each category. A GC was used to find out the amount of adsorption of compounds. Next step was to realize the relation between properties of the compounds and amount of extraction that was carried by ANN.

Experimental

Chemicals and materials

Clomipramine, desipramine, imipramine and amitriptyline were obtained kindly from Arasto Pharmaceutical Chemical Inc. (Tehran, Iran). PAHs were purchased from Merck (Darmstadt, Germany) and used as received. Phthalate esters were obtained from Aldrich (Milwaukee, WI, USA) and hydrofluoric acid was purchased from Fluka (Switzerland). HPLC grade methanol and distilled water were used.

Apparatus

Gas chromatographic analyses of the analytes were performed by an Agilent 7890A GC system (USA), equipped with a split/splitless injector, a flame ionization detector (FID) system and a SB5 column (30 m×0.25 mm I.D., 0.25 μm film thickness)

(SGE, USA). The oven temperature program was different for each group for compounds to separate peaks in short time. The injector and detector temperatures were set at 280 °C and 300 °C, respectively. The mode of injection was splitless. A Parker Hannifin (England) G6010E nitrogen generator and dry air generator were used for the simultaneous nitrogen and air supplying as the carrier and oxidant gases, respectively. A homemade SPME syringe was used for extraction. A homemade thermoelectric cooler was used for cooling of the fiber [5]. Neural Power Ver. 2.5 (CPC-X software, USA) was used to ANN analysis of data. Calculated properties of each compound were obtained from <https://chemicalize.com> (ChemAxon Ltd.).

Preparation of SPME fibers

The steel wires (4 cm in length) were dropped in hydrofluoric acid for 15 min at 40 °C to prepare fibers [10]. The steel wires then rinsed with distilled water several times. The fibers were injected to a GC inlet at temperatures 240 °C and 280 °C and hold 30 min for conditioning. The fibers were inserted into the GC inlet again to check the conditioning. A blank chromatogram was used to confirm that there was no extraneous peak which could be assign to analyte peaks.

SPME procedure

For the fast vaporized solid phase microextraction (FV-SPME) process [5], 500 μL of aqueous solution of the selected analytes were put in 5 mL glass vials, Then the vial was tightly capped with a polytetrafluoroethylene septum and a screw cap. A heater with a glass bath was employed to adjust the temperature of glass bath at 240 $^{\circ}\text{C}$, resulting in instantly evaporation of the sample. A homemade double side thermoelectric cooler was used for reduce temperature of the fibers during as well as a few seconds sooner, extraction [5]. Extractions were carried out by exposing of the fiber to the vial inner room. After extraction, the fiber was withdrawn into a needle, removed from the sample vial and immediately introduced into the GC inlet for the thermal desorption. The extraction time was optimized for each set of analytes.

GC analyzed

Peak area of each analyte was compared to 2-heptanole peak area at a single concentration so a new and normalized value for extractions has been calculated. This extraction value (EV) was used in chemometrics analysis. Different column temperature programs have been used for each category of the analytes.

ANN analysis

Calculated properties and EVs of each

analyte were inserted in the software. 9 analytes were used for training and others three for testing.

The effect of some parameters of ANN include type of algorithm, root square means of error (RSME) and count of neurons, on ANN analysis have been studied.

All data inserted in an excel worksheet. And calculations were carried out there.

Mean absolute percentage error (MAPE, eq. 1) was used for comparison of effects.

$$MAPE = \sum_{i=1}^n \frac{|EV_{prdi} - EV_{expi}|}{EV_{expi}} \times \frac{100}{n} \quad (1)$$

EV_{prdi} and EV_{expi} are EV of predicted and experimental of i th analyte of n analytes, respectively. Because results of each run of ANN are not same, the predicted EVs were average of four calculated EVs.

Results and discussion

The careful choice of network training data and network parameters were essential to succeeding an efficient training process. The physico-chemical properties of a compound maybe affect its adsorption on a single fiber, so it is expected that more properties as the input data will conduct a more solid model. A well-selected training data set is the key factor to receive the best results. The training data must be as possible as close to whole data. Use of too small data set or too specific to a certain portion of overall data set, will reduce accuracy of the model. It is

expected that large training data sets are more similar to whole data set, but obtain and analysis of such large data set are not economical or even possible often. Chromatographic and extraction methods are in this category. Use of large data set are practical for some researches which want shows the ability of the method by results as good as possible, but large data set is not an option for many separation or electrochemistry research. The obtained results are shown a small but well selected sample is enough for a raw, correct and somewhat precise prediction of SPME extraction.

ANN design

The physico-chemical properties of a compound maybe affect its adsorption on a single fiber, so 17 properties are selected and their computational values for each

compound calculated (Table 1). 12 compounds from four categories include polycyclic aromatic hydrocarbons (PAH), phthalate esters, tricyclic antidepressant (TCA) and synthetic opioids were selected (Table 1). These categories are different in structure as much as possible among in ordinary and available analytes which have proper GC analysis properties. Minimum number of the analytes will lead to reduction of analysis time and cost.

For ANN, amitriptyline and two random compounds have been chosen for testing and others 9 for training purpose. No more than one analyte was chosen from each category. The effect of some parameters on results was also studied. The entire process has been illustrated in Figure 1.

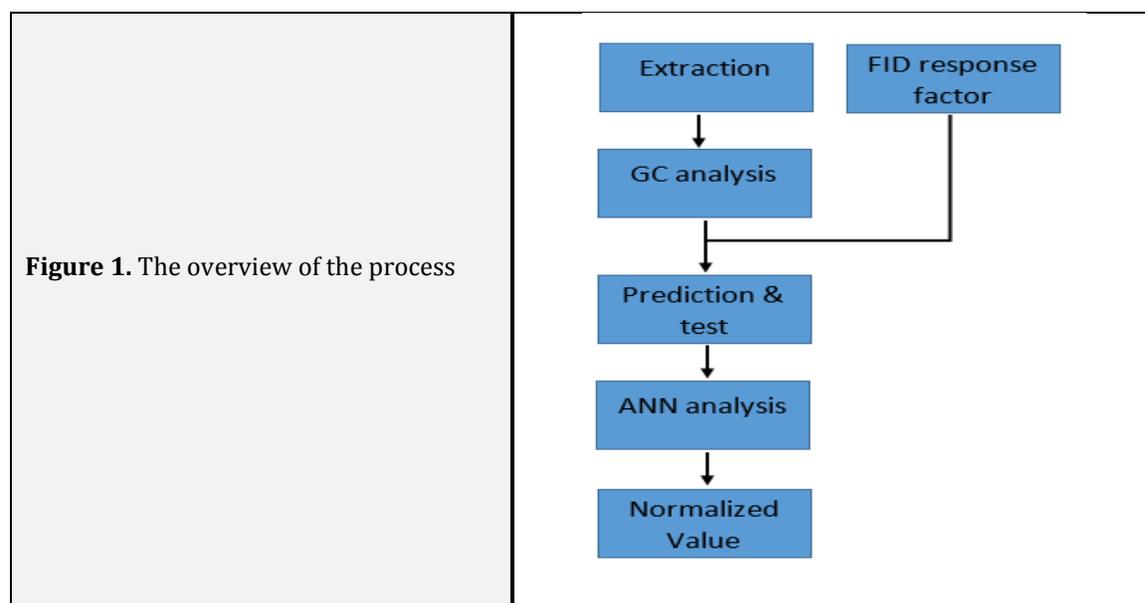


Table 1. Calculated properties of analytes of training (top 9) and testing (bottom 3) of ANN

Analyte	Molar weight	Nitrogen count	Oxygen count	Carbon count	Chlorine count	Hydrogen donor	Hydrogen acceptor	Topological polar surface area	Polarizability	Pk _a	Pk _b	Log P	Log D	Van Der Waals volume	Waals Area	Aromatic ring	Hetro ring
di-Ethyl phthalate	222.24	0	4	12	0	0	2	52.6	22.87	Na ^a	Na	2.69	2.69	206.37	347.68	1	0
i-Butyl phthalate	278.348	0	4	16	0	0	2	52.6	30.22	Na	Na	4.469	4.469	274.44	467.34	1	0
Diocetyl phthalate	390.564	0	4	24	0	0	2	52.6	44.95	Na	Na	8.19	8.19	411.65	714.04	1	0
Naphtalene	128.174	0	0	10	0	0	0	0	18.08	Na	Na	2.96	2.96	125.17	198.52	2	0
Anthracene	178.234	0	0	14	0	0	0	0	25.77	Na	Na	3.95	3.95	168.31	261.04	3	0
Methadon	309.453	1	1	21	0	0	2	20.31	38.17	Na	9.12	5.01	2.904	322.46	540.08	2	0
Desipramine	266.388	2	0	18	0	0	2	15.27	32.82	10.02	0	3.9	1.08	268	441.61	2	1
Clomipramine	314.86	2	0	16	1	0	2	6.48	36.47	9.2	0	4.88	2.72	299.73	494.38	2	1
Imipramine	280.415	2	0	19	0	0	2	6.48	34.67	9.2	0	4.28	2.1	285.75	477.91	2	1
Pyren	202.256	0	0	16	0	0	0	0	30.03	Na	Na	4.28	4.28	184.13	279.09	4	0
Tramadole	263.381	2	1	16	0	1	3	32.7	30.84	13.8	9.23	2.45	0.239	272.29	466.77	1	0
Amitriptyline	277.411	1	0	20	0	0	1	3.24	35.25	9.76	0	4.81	2.12	282.76	457.21	2	0

^a 50000 was used for calculation

Extraction with fiber

Maybe the etched steel fiber is not the most powerful adsorbent among the homemade SPME fibers, but it fits on the purpose of research. This fiber is very simple prepared and has very good stability and durability, or in three words: a simple, good and strong fiber.

FV-SPME [5] is chosen as the mode of extraction. It is very fast and without many optimization parameters. Indeed, at this work, the only parameter was time of extraction. The EVs at optimized extraction condition were used in ANN analysis. GC analysis was just a way to find amount of extraction in presented work. But GC detector response to various analytes are not same and raw FID signals relate to EV and FID response factor. To eliminate the detector response factor, the ratio of GC peak area of direct injection of 2-heptanol to the direct injections of analytes was used. A new value, EV which is depend just to extraction, was calculated in each case (Eq 2).

$$EV = \frac{GC S_{hep}}{GC S_{ana}} \times S_{ana} \quad (2)$$

$GC S_{hep}$ and $GC S_{ana}$ are peaks areas of direct injection of 2-heptanol and analyte to GC, respectively. S_{ana} is peak area of extracted analyte. The chromatograms of direct injection and after SPME extraction of analytes are shown at Figure 2.

Optimization of ANN

The most used approach to find the optimal conditions of ANN is trial and error. The RSME is most used criteria for judgment about the better situation. Unfortunately, this test is weight-dependent and bigger values have more effect or more simply; all analytes are not equal. The absolute relative deviation (ARD) is an alternative. A less common and similar to ARD test (MAPE) which all analytes have same effect, was used for testing. The effect of type of algorithm, RSME and number of nodes were studied.

Type of algorithm

The effect of different algorithm on ANN, bath back propagation [11], incremental back propagation [12], genetic algorithm [13], quickprop [14] and levenberg-marquardt [15] on ANN were studied (Figure 3). Other factors were kept fixed (RSME 50 and Number of neurons 9). The MAPE results of tests are shown in Figure 2. IBP and LM showed worst results (54.8 and 41.8 respectively). The results of QP and GA are close (36.5 and 38.5 respectively). The Best result belongs to BBP with 18.0. BBP will be used in future studies.

Optimization of RSME

RSME is used by Software to find the degree of fitting. When the world and training data set are exactly similar, Less RSME concomitants

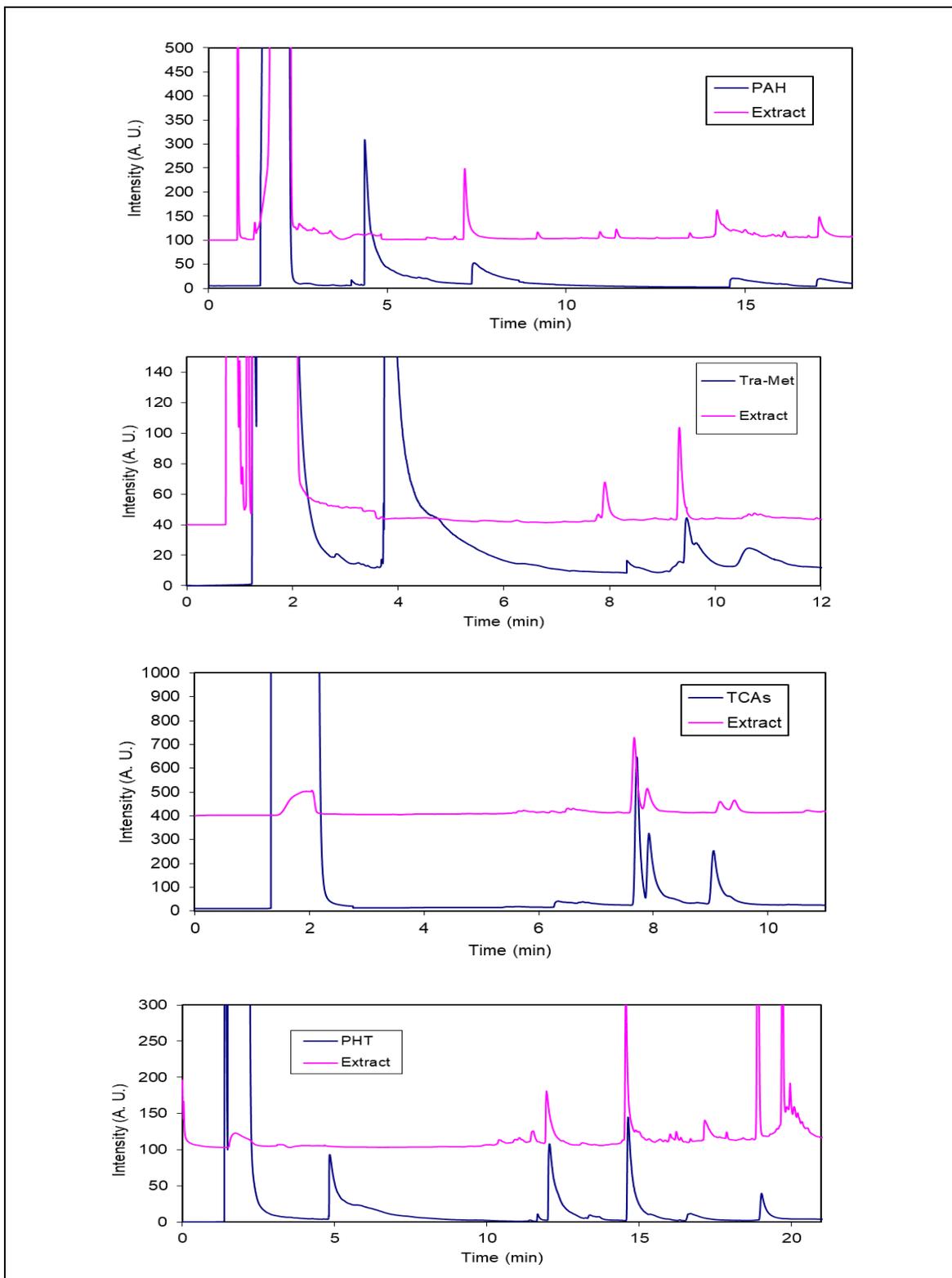


Figure 2. The chromatograms of direct injection and after extraction of analytes, separated to four categories. A peak in about 5 min is for 2-Heptanol

Figure 3. The effect of type of algorithm on MAPE

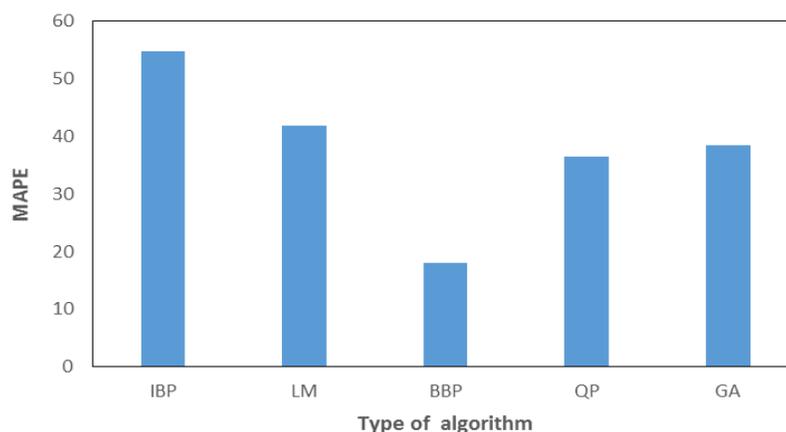
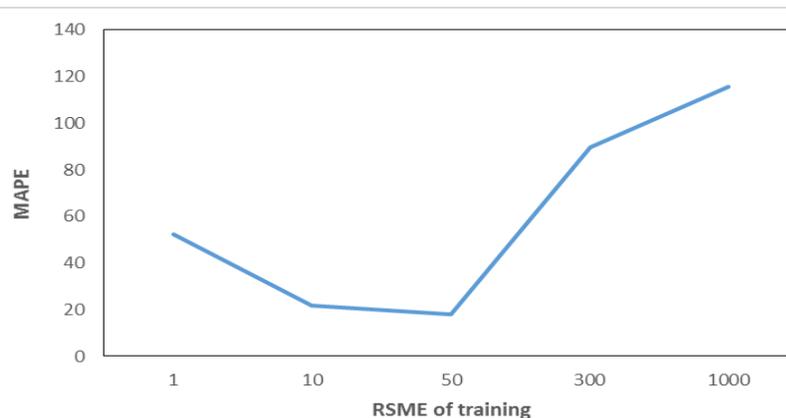


Figure 4. The effect of RSME of training on MAPE



better answer. But a little difference between the world and training set is regular. The over-fitting is a phenomenon which presents in case of very little RSMEs when the fitting model are exactly fitted on training data set, so this fitting is not proper for the world [16].

Overall, very little or very large RSMEs do not lead to the best answer. The effect of RSME of training on MAPE of testing is showed at Figure 4. It is seeming that 50 is a minimum and MAPE increased in both side with increase and decrease of RSME. 50 was

used for future experiments.

Optimization of number of neurons

Neurons are in the hidden layer. They connect input to output. The effect of number of neurons on MAPE is depicted in figure 4. As shown, 5, 7, 9 and 12 lead to 85.1, 29.8, 18 and 30, so the central point (9) is the best and chosen as the optimum (Figure 5).

Characteristic of the method

The experimental and the predicted EVs of training and testing are presented in Table 2

Figure 5. The effect of neuron count on MAPE

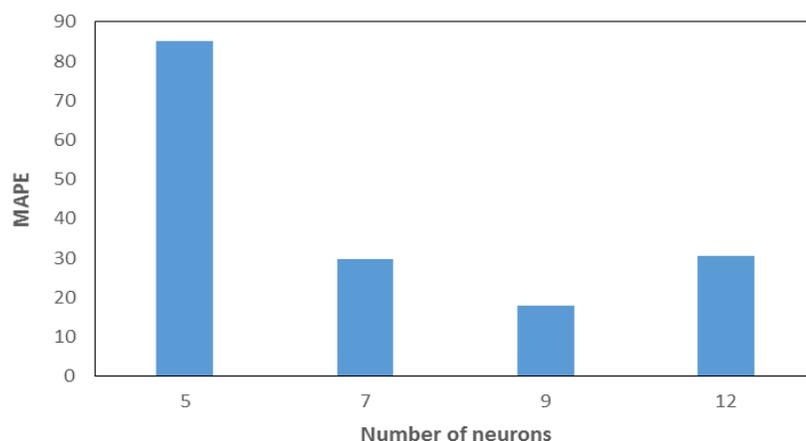
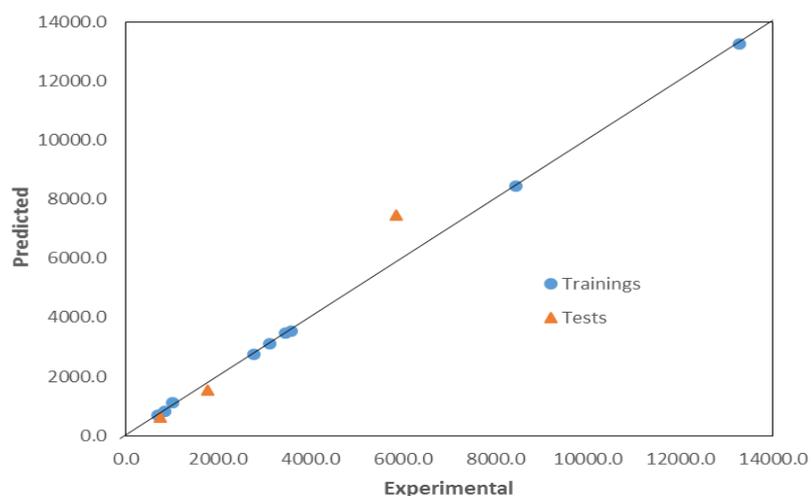


Figure 6. The plot of predicted to experimental values. The guideline is the diagonal line



As obvious, predicted values are in right direction, although there is a little difference. Even more the MAPE of tests are good (18.0%). A difference between training and testing analytes are available that is normal.

The training data set is just a model of the world. Another useful chart is Figure 5, which is predicted to experimental EVs. The distance from diagonal line are an estimation of error of

ANN. It is clear that predictions are very close to experimental data. Large values have more effect in Figure 6. To normalize the values, Figure 7 is presented. The ratio of prediction to experimental EVs is depicted in the latter figure. This figure shows symmetric scattering along experimental value.

ANN, as a learning machine, constitutes of three different layers: the input layer, at least one hidden layer and the output layer (Figure 8).

Figure 7. The predicted to experimental ratio for all analytes include tests and trainings

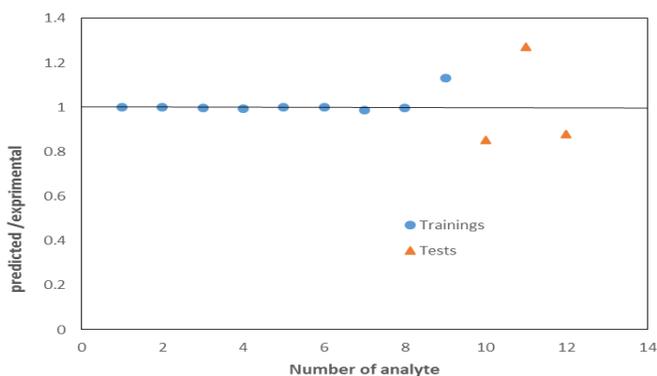


Figure 8. The structure of ANN

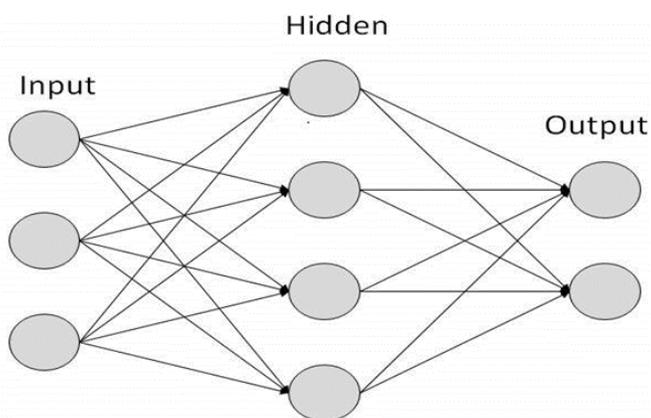


Table 2. Experimental and predicted EVs for validation of ANN model

	Experimental	Calculated	Difference	Percentage of change
di-Ethyl phthalate	3106.0	3108.7	2.7	0.087
i-Butyl phthalate	3451.0	3455.2	4.2	0.12
Diocetyl phthalate	13271.6	13240.8	-30.7	-0.23
Naphtalene	828.9	823.7	-5.2	-0.63
Anthracene	690.2	690.4	0.2	0.029
Methadon	8441.1	8440.6	-0.5	-0.0054
Desipramine	3583.0	3539.4	-43.7	-1.2
Clomipramine	2770.5	2766.9	-3.6	-0.13
Imipramine	1001.6	1131.8	130.2	13
Pyren	743.2	634.4	-108.8	-15
Tramadole	5850.5	7447.6	1597.2	27
Amitriptyline	1761.0	1549.0	-212.0	-12

The hidden layer(s) contain(s) some neurons or nodes that connect input to output layer. First and main step in ANN is training, which

introduce input and output data to program and make a connection between them. The purpose of training is to simulate whole data set (the

world) with all relations and connections by use of a set of small data set (training data set). ANN with a selected algorithm will try to find best weights in the hidden layer(s). For sure of validity of the model, ANN must consider some tests. For this propose, ANN will examine itself with some known data set that uses them as the unknown data set. Several different algorithms are used for ANN. Because ANN answer is not unique, different algorithm should consider to find better solution.

Conclusion

The presented method can be used for prediction of adsorption on a surface coating specifically on a SPME fiber. The method is straightforward, economic, easy and accurate. If the analysis is done with a fast instrument (such as ion mobility spectrometry or florescence spectrometry), it will be rapid, too. By some easy and traditional steps include: extractions, instrumental analysis and finally, a few statistical and computational analysis, the adsorption on the fiber is predicable. Although the predictions are with errors such as 18.0 % (this work) but the estimations are quiet correct, generally. This method can used in two cases: prediction of adsorption of an analyte on different fibers (best fiber) and prediction of adsorption of analytes on a specific fiber (best analytes). Probably, this process is effective for similar

separation methods.

Acknowledgements

The author gratefully acknowledges the Shahid Beheshti University Research Affairs for their support.

References

- [1]. M. Jafari, R. Sedghi, H. Ebrahimzadeh, *Microchim. Acta*, **2016**, *183*, 805-812.
- [2]. M. Jafari, H. Ebrahimzadeh, M.H. Banitaba, S.S. Davarani, *J. Sep. Sci.*, **2014**, *37*, 3142-3149
- [3]. A. Díaz, A. Peña-Alvarez, *J. Chromatographic Sci.*, **2017**, *55*, 946-953.
- [4]. Z. Bai, A. Pilote, P.K. Sarker, G. Vandenberg, *J. Pawliszyn, Anal. Chem.*, **2013**, *85*, 2328-2832.
- [5]. M. Jafari, H. Ebrahimzadeh, M.H. Banitaba, *Talanta*, **2015**, *144*, 474-479.
- [6]. A. Spietelun, A. Kloskowski, W. Chrzanowski, J. Namiesń, *Chem. Rev.*, **2013**, *113*, 1667-1685.
- [7]. M. Buscema, V. Consonni, D. Ballabio, A. Mauri, G. Massini, M. Breda, R. Todeschini, *Chemomet. Intellig. Laboratory*, **2014**, *138*, 110-119.
- [8]. E. Golubović, A. Protić, M. Zečević, B. Otašević, M. Mikić, *J. Chemomet.*, **2014**, *28*, 567-574.
- [9]. M. Khajeh, M. Gharan, *J. Chemomet.*, **2014**, *28*, 539-547.
- [10]. H.L. Xu, Y. Li, D.Q. Jiang, X.P. Yan, *Anal. Chem.*, **2009**, *81*, 4971-4977.

- [11]. B.J. Wythoff, *Chemomet. Intellig. Laboratory*, **1993**, *18*, 115-155.
- [12]. L.M. Fu, H.H. Hsu, J.C. Principe, *IEEE Transact. Neural, Networks*, **1996**, *7*, 757-561.
- [13]. J.H. Holland, *J. Associat. Comput. Machin.*, **1962**, *3*, 297-314.
- [14]. S.E. Fahlmann, C. Lebiere, The cascade-correlation learning architecture. In *Advances in neural information processing systems*, **1990**, pp 524-532.
- [15]. M.T. Hagan, M. Menhaj, *IEEE Transact. Neural Networks*, **1994**, *5*, 989-993.
- [16]. T. Bolanca, S. Cerjan-Stefanovic, S. Ukic, M. Rogosic, M. Luša, *J. Chemomet*, **2008**, *22*, 106-113.

How to cite this manuscript: Mostafa Jafari, The Prediction of Adsorption Properties of a Solid Phase-Microextraction Fiber by Chemometrics methods, *Advanced Journal of Chemistry-Section A (Adv. J. Chem. A)*, **2019**, *2*(3), 203-215.