



J. Serb. Chem. Soc. 76 (12) 1627–1637 (2011)  
JSCS–4235

**Chemical composition of the essential oils of *Citrus sinensis* cv. Valencia and a quantitative structure–retention relationship study for the prediction of retention indices by multiple linear regression**

PARVIZ ABEROOMAND AZAR<sup>1</sup>, MEHDI NEKOEI<sup>1\*</sup>, KAMBIZ LARIJANI<sup>1</sup>  
and SAKINEH BAHRAMINASAB<sup>2</sup>

<sup>1</sup>Department of Chemistry, Science and Research Branch, Islamic Azad University, Tehran  
and <sup>2</sup>Department of Chemistry, Shahrood branch, Islamic Azad University, Shahrood, Iran

(Received 18 December 2010, revised 27 February 2011)

**Abstract:** The chemical composition of the volatile fraction obtained by head-space solid phase microextraction (HS–SPME), single drop microextraction (SDME) and the essential oil obtained by cold-press from the peels of *C. sinensis* cv. *Valencia* were analyzed employing gas chromatography–flame ionization detector (GC–FID) and gas chromatography–mass spectrometry (GC–MS). The main components were limonene (61.34, 68.27 and 90.50 %), myrcene (17.55, 12.35 and 2.50 %), sabinene (6.50, 7.62 and 0.5 %) and  $\alpha$ -pinene (0, 6.65 and 1.4 %) respectively obtained by HS–SPME, SDME and cold-press. Then a quantitative structure–retention relationship (QSRR) study for the prediction of retention indices (*RI*) of the compounds was developed by application of structural descriptors and the multiple linear regression (MLR) method. Principal components analysis was used to select the training set. A simple model with low standard errors and high correlation coefficients was obtained. The results illustrated that linear techniques such as MLR combined with a successful variable selection procedure are capable of generating an efficient QSRR model for prediction of the retention indices of different compounds. This model, with high statistical significance ( $R^2_{\text{train}} = 0.983$ ,  $R^2_{\text{test}} = 0.970$ ,  $Q^2_{\text{LOO}} = 0.962$ ,  $Q^2_{\text{LGO}} = 0.936$ ,  $REP(\%) = 3.00$ ), could be used adequately for the prediction and description of the retention indices of the volatile compounds.

**Keywords:** *Citrus sinensis* cv. *Valencia*; volatile constituents; HS–SPME; SDME; QSRR.

INTRODUCTION

The genus *Citrus* (Rutaceae) is represented in Iran by the species *C. sinensis*, *C. medica*, *C. limon*, *C. nobelis*, *C. aurantifolia* and *C. aurantium*. Citrus fruits

\* Corresponding author. E-mail: m\_nekoei1356@yahoo.com  
doi: 10.2298/JSC101218141A

are the most common subtropical crops in the world. There is a great amount of variation among citrus species and cultivars as a result of frequent bud mutation, interspecific and intergeneric hybridization, apomixis and a long history of cultivation. In Iran, there are many citrus variants the phylogeny of which remains unknown.<sup>1</sup> The essential oils of *Citrus* are placed within the glands in the outer layer of the fruit skin. This oil is composed of many constituents, including monoterpenes, sesquiterpenes, alcohols, esters and aldehydes. The most valuable oils are those of orange and lemon. Cold-press peel oils are generally in use in many food, confectionary, drug, cosmetic and flavoring products.<sup>2</sup> Several studies were performed on the composition of the essential oils from leaves and peel of *C. sinensis* and its hybrids and on their biological activities, such as antifungal, antioxidant and antiaflatoxicogenic.<sup>3–10</sup> Although *C. sinensis* (orange) is one of the most important horticulture products of Iran, a literature survey revealed that there are no reports of an adequate comparative study of the volatile constituents of *C. sinensis* from Iran in which a variety of extraction techniques were employed. Conventional sampling methods for the extraction of *Citrus* essential oil in previous studies were mainly cold-press and hydro-distillation (HD). The cold-press method is mainly used for *Citrus* fruit peel. However, both volatile and non volatile compounds are simultaneously extracted by cold-press and their separation would be inevitable. HD usually requires large amount of samples, long time (several hours), and high energy. Moreover, many unstable aroma volatiles may be thermally decomposed and degraded during thermal extraction or distillation.<sup>11</sup>

Recently, many kinds of extraction techniques, such as single-drop micro extraction (SDME)<sup>12–14</sup> and solid-phase microextraction (SPME),<sup>15–18</sup> have been developed. The SDME technique involves extraction of analytes from a mixture into a microdrop of an organic solvent suspended from the tip of a microsyringe. After extraction, the microdrop is retracted back into the microsyringe and injected into a GC–MS instrument for analysis. The SPME technique is performed using a fused silica fiber that is coated with different stationary phases and is characterized by its high sensitivity to volatile natural compounds. These methods are rapid, simple and inexpensive sample preparation techniques for the extraction and pre-concentration of volatile compounds. Generally, the use of SDME and SPME to extract analytes from a matrix is mainly performed by direct immersion (DI) and headspace (HS). Especially, HS–SPME is considered as a good choice for sample preparation in fragrance and aroma analyses.<sup>19</sup>

The Kovatz retention indices is the key tool for identification of diverse natural compounds present in a volatile oil separated by a variety of isolation techniques. The search for quantitative relationships between molecular structure and retention indices is a basic task in chemistry. Quantitative structure–retention relationship (QSRR) analysis is now a well-established and highly respected te-

chnique to correlate diverse simple and complex physico-chemical properties of a compound with its molecular structure, through a variety of descriptors. The basic strategy of QSRR analysis is to find optimum quantitative relationships, which can then be used for the prediction of the properties from molecular structures. Once a reliable relation has been obtained, it is possible to use it to predict that same property for other structures not yet measured or even not yet prepared.

QSRR for retention indices have been reported for different types of organic compounds.<sup>20–28</sup> The application of these techniques usually requires variable selection for building well-fitted models. In this work, the elimination selection–stepwise regression (ES–SWR) variable selection method was employed. The proposed methodology was validated using several strategies: cross-validation, external validation using division of the entire data set into training and test sets and *Y*-randomization. The aim of this work was the investigation of the chemical composition of *C. sinensis* cv. *Valencia* volatile compounds and also the construction an accurate quantitative relationship between the molecular structure and the retention indices by the stepwise-multiple linear regression (MLR) method.

## EXPERIMENTAL

### *Plant material*

The fruits of *C. sinensis* cv. *Valencia* were collected from Ramsar, Province of Mazandaran, Iran, in January 2008. The material plant was identified and a voucher specimen was deposited at the Herbarium of the *Citrus* Research Institute of Ramsar, Mazandaran, Iran.

### *Isolation of the essential oil*

The essential oil of *C. sinensis* cv. *Valencia* was extracted by cold-press. The volatile constituents were extracted by HS–SPME and HS–SDME. Accordingly, 1.5 g of the peel of *C. sinensis* was placed in a 20 ml vial with screw caps and PTFE/silicone septa. The vial was immersed in a controlled water bath at 70 °C for 30 min. A 100 µm polydimethylsiloxane (PDMS) fiber (Supelco, USA) was used to extract the compounds which evaporated to head-space of the vial (15 min). The analytes were thermally desorbed for 3 min at 250 °C in a splitless GC injector. The equilibration temperature and time were 50 °C and 30 min, respectively. SDME extraction was performed using 2 µl hexadecane (Merck, Germany) as the extraction solvent. The temperature and time conditions were similar to the SPME conditions.

### *Analysis of the volatile compounds*

*Gas chromatography analysis.* Analytical GC was performed on HP-6890 GC system (Hewlett-Packard, USA) equipped with a flame ionization detector (FID) and a HP-5 capillary fused silica column (30 m×0.25 mm ID, film thickness: 0.25 µm). The oven temperature was held at 60 °C for 3 min then programmed at rate of 6 °C min<sup>-1</sup> to 250 °C and held isothermally for 3 min. The carrier gas was nitrogen at a flow rate of 1 ml min<sup>-1</sup>; the split/splitless injector temperature was 250 °C.

*Gas chromatography–mass spectrometry analysis.* The GC–MS analyses were performed on a HP-6890 gas chromatograph (Hewlett-Packard, USA) coupled to an HP-5973 quadrupole mass spectrometer (Hewlett-Packard, USA). The analytes were separated on a HP-5MS capillary column (30 m×0.25 mm with a phase thickness of 0.25 µm). The split/splitless injector temperature was set at 250 °C and the temperature program was 60 °C for 3 min, 6 °C min<sup>-1</sup>

ramp rate to 250 °C and held constant for 3 min. The carrier gas was helium (99.999 %) at a 1 ml min<sup>-1</sup> flow rate. In the SPME analysis, splitless injection (3 min) was used at 250 °C. The mass spectrometer was operated in the electron-impact mode (EI) at 70 eV.

#### *Qualitative and quantitative analysis of the compounds*

The compounds in each sample were identified by comparison of their mass spectral pattern and their linear retention indices based on a homologous series of even normal alkanes (C<sub>8</sub>–C<sub>24</sub>) with those of authentic references<sup>29</sup> and the Wiley 257 mass spectra database. The percentage of each compound was calculated from peak area obtained by FID.

#### *QSRR study*

*Calculation of molecular descriptors.* Molecular descriptors are defined as numerical characteristics associated with chemical structures. A molecular descriptor is the final result of a logic and mathematical procedure which transforms chemical information encoded within a symbolic representation of a molecule into a useful number, which is applied to correlate physical properties. The Dragon software (Milano Chemometrics and QSAR Research Group, Milan, Italy) was used to calculate the descriptors in this study and a total of 1481 molecular descriptors were calculated for each molecule. Since the values of many descriptors are related to the bond lengths and bond angles *etc.*, the chemical structure of every molecule must be optimized before calculation of its molecular descriptors. For this reason, the chemical structures of the 25 studied molecules were drawn with Hyperchem software (version 7.0 Hypercube, Alberta, Canada) and saved with the HIN extension. To optimize the geometry of these molecules, the AM1 geometrical optimization was applied. After optimization of the chemical structures of all compounds, the molecular descriptors were calculated using Dragon. A wide variety of descriptors have been reported in the literature, having been used in QSRR analyses.<sup>30-35</sup>

*Stepwise multiple linear regression.* As described in the introduction section, the ES–SWR algorithm<sup>30</sup> was used to select the most appropriate descriptors. ES–SWR is a popular stepwise technique that combines forward selection (FS–SWR) and backward elimination (BE–SWR). It is essentially a forward selection approach, but at each step it considers the possibility of deleting a variable as in the backward elimination approach, provided that the number of model variables is greater than two.

*Model validation.* The stability and robustness of the proposed MLR model was illustrated using the following evaluation techniques: leave-one-out (LOO) and leave-group-out (LGO) cross-validation procedures, validation through an external test set and *Y*-randomization.

Cross-validation is a popular technique used to explore the reliability of statistical models. Based on this technique, a number of modified data sets are created by deleting in each case one or a small group (leave-group-out) of objects. For each data set, an input–output model is developed, based on the utilized modeling technique. The model is evaluated by measuring its accuracy in the prediction of the responses of the remaining data stands that were not utilized in the development of the model.<sup>36</sup>

The *Y*-randomization technique ensures the robustness of a QSRR model. The dependent variable vector (*RI*) is randomly shuffled and a new QSRR model is developed using the original independent variable matrix. The new QSRR models (after several repetitions) are expected to have low *R*<sup>2</sup> and *Q*<sup>2</sup> values. If this is not the case, then an acceptable QSRR model cannot be obtained for the specific modeling method and data.

## RESULTS AND DISCUSSION

*Chemical composition of the essential oil*

The volatile components of fruit peel of *C. sinensis* cv. *Valencia* that were isolated by SPME, SDME and cold-press are listed in Table I. Thirteen compounds (representing 99.77 %), ten compounds (representing 98.46 %) and fourteen compounds (representing 97.9 %) were experimentally identified, respectively. The oils were rich in monoterpenes: limonene (61.34, 68.27 and 90.5 %), as the major component, followed by myrcene (17.55, 12.35 and 2.50 %); sabinene (6.50, 7.62 and 0.5 %);  $\alpha$ -pinene (0.0, 6.65 and 1.4 %) in the oils obtained by the SPME, SDME and cold-press methods, respectively.

TABLE I. Chemical composition of the essential oil of *C. sinensis* cv. *Valencia* isolated by the SPME, SDME and cold-press methods and the corresponding observed and predicted *RI* values by SW-MLR for the training and test set

Compound	SPME, %	SDME, %	Cold-press, %	<i>RI</i> (Exp.)	<i>RI</i> <sup>a</sup> (Pred.)	<i>E</i> <sup>b</sup> / %
<b>1a</b> Octane	8.37	0.1	–	800	816.12	2.01
<b>2a</b> 4-Methylthiazole	–	0.14	–	819	856.29	4.55
<b>3b</b> <i>n</i> -Nonane	–	–	0.2	900	953.18	5.90
<b>4a</b> $\alpha$ -Pinene	–	6.65	1.4	939	932.38	–0.70
<b>5a</b> Sabinene	6.50	7.62	0.5	975	1010.11	3.60
<b>6b</b> Myrcene	17.55	12.35	2.5	991	1047.60	5.71
<b>7a</b> <i>n</i> -Octanal	–	0.19	0.8	999	1001.93	0.29
<b>8a</b> $\alpha$ -Phellandrene	–	0.33	–	1003	1033.04	2.99
<b>9a</b> <i>iso</i> -Sylvesteren	1.14	–	–	1009	1016.00	0.69
<b>10b</b> Limonene	61.34	68.27	90.5	1029	1040.46	1.01
<b>11a</b> <i>E</i> - $\beta$ -Ocimene	–	0.5	0.1	1050	1025.95	–2.28
<b>12a</b> $\gamma$ -Terpinene	0.5	–	0.2	1060	1053.73	–0.59
<b>13b</b> Terpinolene	–	–	–	1089	1046.97	–3.85
<b>14a</b> Linalool	1.84	–	0.8	1097	1089.45	–0.68
<b>15a</b> $\alpha$ -Thujene	–	2.31	–	1114	1067.35	–4.18
<b>16b</b> <i>trans</i> -Limonene oxide	0.18	–	–	1142	1160.10	1.58
<b>17a</b> Citronellal	0.29	–	0.2	1153	1146.40	–0.57
<b>18a</b> Decanal	1.40	–	–	1202	1158.00	–3.65
<b>19b</b> $\beta$ -Elemene	0.08	–	0.1	1391	1382.56	–0.60
<b>20a</b> Tetradecane	–	–	0.2	1400	1360.02	–2.85
<b>21a</b> $\beta$ -Caryophyllene	0.1	–	–	1419	1419.46	0.03
<b>22a</b> <i>trans</i> -Uurolo-3,5-diene	–	–	–	1454	1461.25	0.49
<b>23b</b> Valencene	0.48	–	–	1496	1493.75	–0.15
<b>24a</b> Pentadecane	–	–	0.2	1500	1559.54	3.96
<b>25a</b> Hexadecane	–	–	0.2	1600	1600.16	0.00
Total	99.77	98.46	97.9			

<sup>a</sup>Predicted by SW-MLR method; <sup>b</sup>relative error

### Comparison of the chemical profile with similar reports

An investigation in China revealed that limonene was observed as the dominant constituent (77.49 %) in the peel oil of sweet orange, followed by myrcene (6.27 %),  $\alpha$ -farnesene (3.64 %) and  $\gamma$ -terpinene (3.34 %).<sup>4</sup> The main compounds in *C. sinensis* from Uganda and Rwanda were limonene (87.9 and 92.5 %), myrcene (2.4 and 2.0 %),  $\alpha$ -pinene (0.5 and 2.4 %) and linalool (1.2 and 0.9 %).<sup>5</sup> Limonene (90.16 and 77.34 %) was the main compound in fresh and dried pericarps of *C. sinensis* in China. Monoterpene hydrocarbons were the most abundant fraction in the oils of three Kenyan *C. sinensis* varieties, *i.e.*, Salustiana (96.9 %), Valencia (94.5 %) and Washington navel (92.7 %) oils. In each oil, limonene,  $\alpha$ -pinene, sabinene and  $\alpha$ -terpinene were the major compounds.<sup>9</sup> The oil of Italian *Citrus sinensis* (L.) Osbeck cv. *Maltese* was characterized by limonene (92.6 %) as the major constituent.<sup>10</sup>

### QSRR results

Principal components analysis (PCA) was performed with the calculated structure descriptors for the whole data set to detect the homogeneities in the data set and to show the spatial location of the samples to assist the separation of the data into training and test sets. The PCA results showed that two principal components (PC1 and PC2) described 69.67 % of the overall variables, as follows: PC1 47.40 % and PC2 22.27 %. As almost all the variables can be accounted for by the first two PCs, their score plot is a reliable representation of the spatial distribution of the points for the data set. The plot of PC1 against PC2 (Fig. 1) displays the distribution of the compounds over the space of the first two principal components.

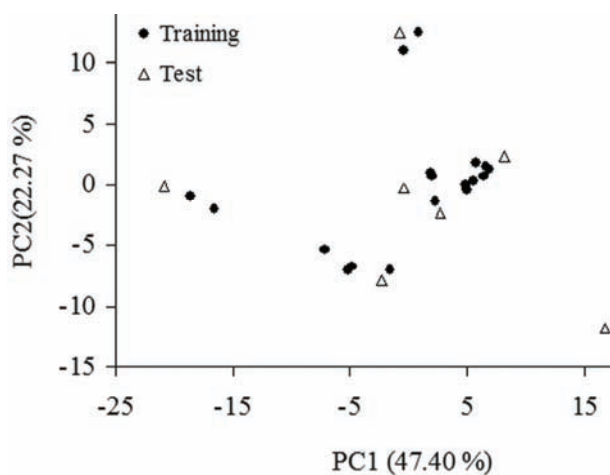


Fig. 1. Principal components analysis of the training and test sets.



According to the results of PCA, all the data were divided into a training set of 18 compounds to develop the models and a test set of 7 compounds to evaluate the models based on two rules:

1. the range of the *RI* values of both the training set and the test set should be covered from the lowest to the highest;
2. the points corresponding to the training set in the PCA plot should not be outside the main clusters. The two sets are listed in Table I.

After analysis of the division of the data set into the training and the test sets, MLR analysis was performed to derive the best QSRR model. The MLR technique was performed on the molecules of the training set given in Table I. After regression analysis, a few suitable models were obtained among which the best model was selected and presented in Eq. (1). A small number of molecular descriptors were used to establish the QSRR model. Additional validation was performed on the external data set (test set) consisting of 7 essential oil compounds. MLR analysis provided a useful equation that could be used to predict the *RI* of an essential oil compound based on these parameters. The result of this study was the development of a new linear QSRR model containing 4 variables. The best equation obtained for the *RI* of the essential oil compounds is:

$$RI = 758.10 + 5.23(Mor01m) - 155.64(Gu) - 895.46(Mor30v) - 246.37(Mor29u) \quad (1)$$

$$N_{\text{training}} = 18; R^2_{\text{training}} = 0.983; RMSE_{\text{training}} = 27.801; \%REP = 3.00; Q^2_{\text{LOO}} = 0.962; Q^2_{\text{LGO}} = 0.936, R^2_{\text{test}} = 0.970; RMSE_{\text{test}} = 34.505$$

In Eq. (1), *N* is the number of compounds in the training set,  $R^2$  is the squared correlation coefficient,  $Q^2_{\text{LOO}}$ ;  $Q^2_{\text{LGO}}$  are the squared cross-validation coefficients for LOO and LGO, respectively, *REP* is the relative error of prediction and *RMSE* is the root mean square error of prediction.

From Eq. (1), it can be concluded that the three of most significant descriptors according to the ES-SWR algorithm are 3D-MoRSE descriptors. Furthermore, the other one belongs to the WHIM descriptors. A brief explanation of the descriptors that were selected is given below.

Three-dimensional MoRSE descriptors are derived from infrared spectra simulation using a generalized scattering function.<sup>30</sup> The three descriptors, appearing in the model are *Mor01m*, *Mor30v* and *Mor29u*. *Mor01m* was proposed as signal 01 / weighted by atomic mass, which relates to the mass of the molecule. *Mor30v* was proposed as signal 30 / weighted by atomic van der Waals volumes, which relates to atomic van der Waals volumes. *Mor29u* was proposed as signal 29 / unweighted.

*Mor01m* displays a positive sign, which indicates that the atomic mass of a molecule is directly related to its retention index. *Mor30v* and *Mor29u* display a

negative sign which indicates that the volume of a molecule is inversely related to the retention index.

*Gu* (*G* total symmetry index / unweighted) is the fourth descriptor appearing in the model. It is one of the WHIM descriptors, which are based on the statistical indices calculated on the projections of atoms along the principal axes. The algorithm consists of performing a principal components analysis on the centered Cartesian coordinates of a molecule using a weighted covariance matrix obtained from different weighing schemes for the atoms. Directional WHIM symmetry descriptors are related to the number of central symmetric atoms (along the  $m^{\text{th}}$  component), the number of asymmetric atoms and the total number of atoms in the molecule. *Gu* displays a negative sign, which indicates that the retention index is inversely related to the *Gu* descriptor.

From the above discussion, it was concluded that the atomic masses and atomic van der Waals volumes are the main independent factors contributing to the retention index of the components of the studied essential oils.

As can be seen from the correlation matrix (Table II), there was no significant correlation between the selected descriptors.

TABLE II. Correlation matrix for the seven selected descriptors

	<i>Mor01m</i>	<i>Gu</i>	<i>Mor30v</i>	<i>Mor29u</i>
<i>Mor01m</i>	1			
<i>Gu</i>	0.285	1		
<i>Mor30v</i>	0.208	-0.411	1	
<i>Mor29u</i>	-0.694	-0.026	-0.625	1

Equation (1) was used to predict the *RI* for the test set. The data set and the corresponding experimental and predicted *RI* values of all the molecules studied in this work are summarized in Table I. A plot of the values predicted by the SW-MLR against the experimental values of the retention indices of the training and test sets is shown in Fig. 2.

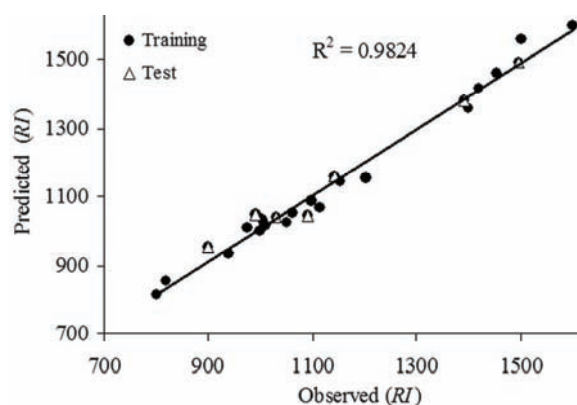


Fig. 2. The *RI* values predicted by MLR modeling vs. the experimental *RI* values.



The residuals (observed  $RI$  – predicted  $RI$ ) vs. the observed  $RI$  value, obtained by the SW–MLR modeling, are shown in Fig. 3. The distribution of the residuals on both sides of the zero line indicates that there is no systematic error in the SW–MLR model.

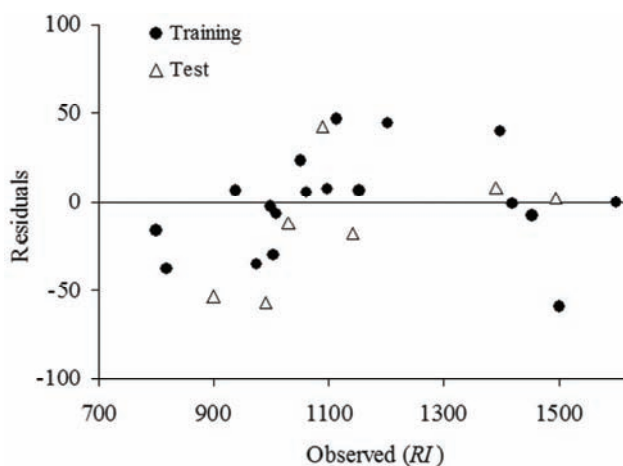


Fig. 3. Plot of the residuals against the experimental values of the retention indices.

The results illustrated once more that the linear MLR technique combined with a successful variable selection procedure is adequate to generate an efficient QSRR model for predicting the  $RI$  values of essential oil components.

For a more exhaustive testing of the predictive power of the model, validation of the model was also performed using the LOO and the LGO cross-validation techniques on the training set of compounds. For the LOO cross-validation, a data point is removed from the set, and the model is recalculated. The predicted  $RI$  for that point is then compared with its actual value. This is repeated until each data point has been omitted once. For LGO, 20 % of the data points are removed from the dataset and the model was refitted; the predicted values for those points were then compared with the experimental values. Again, this is repeated until each data point has been omitted once. The results produced by the LOO ( $Q^2_{LOO} = 0.962$ ) and the LGO ( $Q^2_{LGO} = 0.936$ ) cross-validation tests illustrated the quality of the obtained model.

The model was further validated by applying  $Y$ -randomization. Several random shuffles of the  $Y$  vector ( $RI$ ) were performed and the low  $R^2$  and  $Q^2$  values that were obtained showed that the good results in the original model use were not due to a chance correlation or structural dependency of the training set. The results of the  $Y$ -randomization test are presented in Table III.

The proposed method, due to the high predictive ability and simplicity could be a useful aid to the costly and time consuming experiments for determining the  $RI$  values of the components of essential oils.

TABLE III.  $R^2$  and  $Q^2$  values after several  $Y$ -randomization tests

Iteration	$R^2$	$Q^2$
1	0.082	0.230
2	0.104	0.075
3	0.332	0.061
4	0.072	0.123
5	0.073	0.135
6	0.305	0.044
7	0.182	0.001
8	0.142	0.023
9	0.168	0.009
10	0.241	0.026

## CONCLUSIONS

This paper deals with the characterization of the volatile oils obtained from the peels of *C. sinensis* cv. *Valencia* by the HS-SPME, SDME and cold-press methods. In the all identified oils, monoterpenes dominated over non-terpenes and sesquiterpenes. The QSRR method using stepwise-MLR analysis employed to develop a model for predicting the retention indices of the components of essential oils. The developed QSRR model with simply calculated molecular descriptors could be employed to estimate the retention index for new compounds, even in the absence of standard candidates.

## ИЗВОД

ХЕМИЈСКИ САСТАВ ЕТАРСКОГ УЉА *Citrus sinensis* cv. *Valencia* И QSRR АНАЛИЗА У ПРОЦЕНИ РЕТЕНЦИОНИХ ИНДЕКСА ПРИМЕНОМ МЕТОДЕ ВИШЕСТРУКЕ ЛИНЕАРНЕ РЕГРЕСИЈЕ

PARVIZ ABEROOMAND AZAR, MEHDI NEKOEI, KAMBIZ LARIJANI и AZAM VAFAEI

*Department of Chemistry, Faculty of Basic Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran*

Хемијски састав испарљивих фракција добијених микроекстракцијом у чврстој фази (HS-SPME), микроекстракцијом из капи (SDME), као и састав етарског уља добијеног хладним цеђењем из *C. sinensis* cv. *Valencia* је испитан методама GC-FID и GC-MS. Главни састојци су били лимонен (61,34, 68,27 и 90,50 %), мирцен (17,55, 12,35 и 2,50 %), сабинен (6,50, 7,62 и 0,50 %) и  $\alpha$ -пинен (0,00, 6,65 и 1,40 %). QSRR испитивање за процену ретенционих индекса једињења ( $RI$ ) је развијено користећи методу вишеструке линеарне регресије (MLR). Анализа главних састојака је коришћена у избору модела са ниском стандардном грешком и високим коефицијентом корелације. Резултати су показали да линеарне технике, као што је MLR, у комбинацији са успешним избором променљиве могу успоставити ефикасан QSRR модел за предвиђање  $RI$  различитих једињења. Овај модел се, са високом статистичком значајношћу ( $R^2_{\text{train}} = 0,983$ ,  $R^2_{\text{test}} = 0,970$ ,  $Q^2_{\text{LOO}} = 0,962$ ,  $Q^2_{\text{LGO}} = 0,936$ ,  $REP(\%) = 3,00$ ), може користити за успешно предвиђање и описивање ретенционих индекса испарљивих супстанци.

(Примљено 18. децембра 2010, ревидирано 27. фебруара 2011)

## REFERENCES

1. V. Mozaffarian, *A dictionary of Iranian plants names*, Farhang Mo'aser Publishers, Tehran, Iran, 1998
2. S. Nagy, P. Shaw, M. Veldhuis, *Citrus Sci. Technol.* **2** (1977) 290
3. P. Singh, R. Shukla, B. Prakash, A. Kumar, S. Singh, P. K. Mishra, N. K. Dubey, *Food Chem. Toxicol.* **48** (2010) 1734
4. N. G. To, Y. J. Liu, M. L. Zhang, *Int. J. Food Sci. Technol.* **44** (2009) 1281
5. S. M. Njoroge, N. T. Phi, M. Sawamura, *J. Essential Oil Bearing Plants* **12** (2009) 26
6. M. Viuda-Martos, Y. Ruiz-Navajas, J. Fernández-López, J. Pérez-Álvarez, *Food Control* **19** (2008) 1130
7. C. P. Araújo-Junior, C. A. G. Dacamara, N. C. Ribeiro, C. A. Gomes, M. M. de Moraes, P. S. Botelho, *Nat. Prod. Commun.* **5** (2010) 471
8. N. C. Ribeiro, C. A. G. Dacamara, F. S. Born, H. A. A. Siqueira, *Nat. Prod. Commun.* **5** (2010) 1819
9. S. M. Njoroge, H. Koaze, P.N. Karanja, M. Sawamura, *Flavour Fragr. J.* **20** (2005) 80
10. A. Trozzi, A. Verzera, G. Lamonica, *J. Essent. Oil Res.* **11** (1999) 482
11. A. C. Kimbaris, N. G. Siatis, D. J. Daferera, *Ultrason. Sonochem.* **13** (2006) 54
12. C. H. Deng, Y. Mao, F. L. Hu, X. M. Zhang, *J. Chromatogr., A* **1152** (2007) 193
13. L. Vidal, A. Canals, N. Kalogerakis, E. Psillakis, *J. Chromatogr., A* **1089** (2005) 25
14. D. A. Lambropoulou, E. Psillakis, T. A. Albanis, N. Kalogerakis, *Anal. Chim. Acta* **516** (2004) 205
15. C. H. Deng, X. Q. Xu, N. Yao, N. Li, X. M. Zhang, *Anal. Chim. Acta* **556** (2006) 289
16. Q. Ye, *J. Chromatogr., A* **1213** (2008) 239
17. G. Ouyang, J. Pawliszyn, *Anal. Chim. Acta* **627** (2008) 184
18. N. Li, C. H. Deng, Y. Li, H. Ye, X. M. Zhang, *J. Chromatogr., A* **1133** (2006) 29
19. F. Augusto, A. L. Lopes, C. A. Zini, *Trends Anal. Chem.* **22** (2003) 160
20. C. Ji, Y. Li, L. Su, X. Zhang, X. Chen, *J. Sep. Sci.* **32** (2009) 3967
21. M. H. Fatemi, E. Baher, M. Ghorbanzadeh, *J. Sep. Sci.* **32** (2009) 4133
22. L. Komsta, *Anal. Chim. Acta* **593** (2007) 224
23. B. D. Junkes, R. D. Amboni, R. A. Yunes, V. E. F. Heinzen, *Anal. Chim. Acta* **477** (2003) 29
24. M. Jalali-Heravi, M. H. Fatemi, *J. Chromatogr., A* **915** (2001) 177
25. Z. Garakani-Nejad, M. Karlovits, W. Demuth, T. Stimpfl, W. Vycudilik, M. Jalali-Heravi, K. Varmuza, *J. Chromatogr., A* **1028** (2004) 287
26. J. Acevedo-Martinez, J. C. Escalona-Arranz, A. Villar-Rojas, F. Tellez-Palmero, R. Perez-Roses, L. Gonzalez, R. Carrasco-Velaz, *J. Chromatogr., A* **1102** (2006) 238
27. P. Tulasamma, K. S. Reddy, *J. Mol. Graphics Modell.* **25** (2006) 507
28. K. Heberger, T. Kowalska, *Chemom. Intell. Lab. Syst.* **47** (1999) 205
29. R. P. Adams, *Identification of essential oil components by gas chromatography / quadrupole mass spectroscopy*, Allured Publishing Corp., Carol Stream, IL, USA, 2004
30. R. Todeschini, V. Consonni, *Handbook of molecular descriptors*. Wiley-VCH, Weinheim, Germany, 2000
31. L. B. Kier, L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, RSP-Wiley, Chichester, UK, 1986
32. E. V. Konstantinova, *J. Chem. Inf. Comp. Sci.* **36** (1997) 54
33. G. Rucker, C. Rucker, *J. Chem. Inf. Comp. Sci.* **33** (1993) 683
34. J. Galvez, R. Garcia, M. T. Salabert, *J. Chem. Inf. Comp. Sci.* **34** (1994) 520
35. P. Broto, G. Moreau, C. Vandicke, *J. Med. Chem.* **19** (1984) 66
36. B. Efron, *J. Am. Stat. Assoc.* **78** (1983) 316.