Calibration of mass selective detector in non-target analysis of volatile organic compounds in the air

Gregor Arh [a,*], Leo Klasinc [b], Marjan Veber [a], Matevž Pompe [a]

[a] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, 1000 Ljubljana, Slovenia
[b] Ruđer Bošković Institute, Bijenička 54, HR-10002 Zagreb, Croatia

ABSTRACT

Volatile organic compounds (VOCs) play an important role in the chemistry of the atmosphere and in biogeochemistry. They contribute to the oxidative capacity of the atmosphere, particle and air pollutants, as well as to the production of greenhouse gasses (for instance ozone). Among analytical techniques for their determination in the atmosphere gas chromatography coupled with mass spectrometry (GC–MS) offers several advantages. However, for an accurate quantification calibration with standard substances is necessary. A quantitative structure–property relationship (QSPR) model for the prediction of MS response factors was developed on basis of our experimental measurements for the quantification of ozone precursors present in the atmosphere. A linear correlation between chemical structures and response factors was established by using a 7-parameter MLR model. The average error in the prediction of response factors was calculated by cross-validation procedure and was below 20%, which is sufficient for the determination of VOCs in the air. The proposed procedure is time consuming so it is more suited for the quantification of tentatively identified organic compounds during the reprocessing of MS chromatograms in cases when the original sample is no longer available.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

In recent years research related to the potentially toxic compounds in the environment has become very important. Volatile organic compounds (VOCs) present a threat in the atmosphere, since they act as precursors in photochemical generation of ozone [1]. Ground ozone is a major problem in rural and as well as urban areas because of its adverse impact on human health and on crops and forest ecosystems [2,3]. VOCs are a large group of compounds characterized by the vapor pressure higher than 1.4 hPa under ambient conditions. This group includes alkanes, alkenes, alkylbenzenes, carbonyls, alcohols, carboxylic acids, esters, ethers and chlorinated aliphatic hydrocarbons [4,5]. Exposure to VOCs can cause different health effects, such as acute and chronic respiratory effects, neurological disorders, cancer and eye and throat irritation [6,7].

There are several analytical procedures available for the determination of VOCs in the air, which differ mostly on sampling procedures. VOCs can either be sampled directly into a gas chromatograph or to stainless steel or Teflon cylinders or alternatively they can be pumped through sampling tubes, filled with sorbent materials (carbotrap, carbosieve, tenax, etc.) [8–14]. Most

frequently the sampling is followed by desorption of VOCs into a cryotrap, and separation on chromatographic column. Afterwards the GC detection is usually carried out by either flame ionization detector (FID) or mass selective detector [15–19]. When characterizing environmental samples one most often deals with a non-target analysis, where the type and number of present compounds are unknown. The calibration procedures therefore become very demanding and time consuming or sometimes even impossible. To solve this problem Katritzky et al. [20] proposed a quantitative structure–property relationship (QSPR) treatment in order to predict gas chromatographic retention times and FID response factors. Response factor (RF) is defined as a quotient between peak area counts and quantity of injected compound in ng. The idea of QSPR principles is to establish a correlation between chemical structure and a chosen property. It is based on two main steps: firstly, the chemical compounds are translated into a computer readable form and secondly, a quantitative correlation between chemical structure and chosen property (for instance response factor) is established. For that several different statistical and learning processes, such as multiple linear regression (MLR), artificial neural networks (ANN), partial least-squares (PLS) method, can be used [21–23].

To predict response factors of FID in cases of non-availability of pure specimens, the "effective carbon number" method (ECN) was proposed by Sternberg et al. [24]. This concept was further investigated by many other researchers [25–28]. In general the RF of a

* Corresponding author. Tel.: +386 1 2419 172; fax: +386 1 2419 220.
E-mail address: gregor.arh@fkkt.uni-lj.si (G. Arh).

flame ionization detector is equivalent to the effective number of carbon atoms. For a hydrocarbon the ECN is simply the number of carbon atoms present. Different functional groups contribute to the ECN of a specific compound. For instance aliphatic and aromatic carbon atoms have ECN contribution 1, whereas olefinic has 0.95, acetylenic 1.30, carbonyl as well as carboxyl 0 and nitrile 0.3. Oxygene atom in ethers has contribution −1.0, in primary alcohols −0.5, in secondary alcohols −0.75 and in tertiary alcohols −0.25. More detailed explanation can be found in the literature [24].

However, in the previously mentioned study by Katritzky et al. [20] the authors imply that for the classes of compounds for which the ECN can be calculated, both calculated and measured RF's can in some cases vary even by 25%. For example the uncertainty of determination of ECN for hydrocarbons is less than 10%, while for some heteroatom containing compounds it can be up to 25%. That's the reason the authors developed a QSPR model for the prediction of RFs and as well as GC retention times. Their results demonstrated that GC-FID retention times and RFs can be predicted with a considerable degree of confidence. Despite mentioned weaknesses of the ECN concept, it has been widely used and good prediction results were published in various studies [25–28].

Further studies for the prediction of RFs for FID and some other detectors were also published. Jalali-Heravi and Fatemi [29] used experimental data set from study of Katritzky et al. [20] and successfully developed an ANN for the modeling of FID RFs. In their further work they also developed a QSPR model by using an ANN for the prediction of RFs of the thermal conductivity detector [30]. In a published paper by Jalali-Heravi et al. [31] the prediction of relative RFs for the electron-capture detector for some polychlorinated biphenyls is described. Only recently a QSPR model that enables calculation of MS response factors from the molecular structure was proposed [32], however it was developed just for the hydrocarbons where no heteroatoms are present. The proposed idea for creating a QSPR model for the calibration of MS detector was extended to VOCs containing heteroatoms. The prediction model was created based on our own experimental measurements. Besides thorough model validation using leave-one-out cross-validation procedure, an experimental validation of the applicability of the developed model was performed. An extensive testing of the MS detector stability and tuning mode was done in order to ensure that once the model is created it can be used for a longer period of time.

1.1. Creation of the models

Initially experimental response factors were obtained for 52 different compounds. Afterwards all structures were constructed with the HyperChem software, where also simple geometric optimization was performed. Then the BABLE software was used to translate all HyperChem output files to the MOPAC input files. The MOPAC software package was used for calculation of optimized structural co-ordinates and net atomic charges. AM1 semi-empirical method was used for the 3D geometrical optimization and the calculations of electrostatic potentials. The CODESSA software was used to calculate all possible descriptors (topological, geometric, informational, electrostatic, electrotopological and quantum-chemical) from the MOPAC output files. This software was also used for choosing the best subset of structural descriptors by minimizing errors in prediction using MLR model. The algorithm of this program is to search for the best MLR model with a selected number of n parameters by omitting all information-less descriptors, i.e. descriptors with no variation between structures, descriptors that do not cover the whole modeling space and descriptors with the squared correlation coefficient (r^2) smaller than 0.01. In the next step all squared pairwise correlation coefficients were calculated and one of the descriptors with square pairwise correlation coefficients above

0.9 was removed in order to eliminate collinear descriptors. The inclusion of collinear descriptors in the same model leads to over-training. More thorough description can be found in the literature [33–37].

1.2. Data set and model validation

Our experimental data set contained 52 different alkanes, alkenes, chlorinated hydrocarbons and hydrocarbons with heteroatoms. Compounds in the data set were chosen since they are typically determined when studying the photochemistry of ozone. Since in our case a relative small data set was used, the leave-one-out cross-validation procedure was used to evaluate prediction capabilities of MLR models during stepwise selection of structural descriptors. The cross-validation method is based on prediction of the property value for one compound from the data set (validation data), while the property values of remaining compounds are used as a training data. The process is repeated such that each individual property value is used once as a validation data. The MLR model with the best cross-validation results was chosen for the prediction model. The data set is presented in Table 1.

1.3. Measurements of RFs

The experimental data were obtained by two gas chromatographs, Varian Star 3400 Cx and Varian Star 3600 Cx. The first one was used in conjunction with a flame ionizing detector and the other was coupled with the Saturn 2000 MS detector. Both systems were equipped with a 10-way VICI Valve (Valco Instruments Co. Inc.) and a cryotrap [38]. Samples were injected into a cryotrap, which was cooled with liquid nitrogen (−196 °C), by using a Helium 6.0 carrier gas. All connection tubes were made out of a stainless steel and were heated to approximately 100 °C in order to prevent compounds from liquefying already in the analytical instrument. For the separation of compounds Restek RTX-5MS column (l = 60 m, $2r$ = 250 μm, d = 5 μm) was used. Temperature program was as followed: initial temperature 3 °C (hold time 10 min), temperature gradient 2 °C/min to 140 °C and 20 °C/min to 250 °C (hold time 10 min).

Two multicomponent standard mixtures, Restek VOC AB-18475 and Matheson Toxi-Mat TO-14 VOC, were used. FID was calibrated to a mixture of C1-C6 n-Paraffins (Fluka 80311). Gas standards of hydrocarbons containing heteroatoms were prepared by injecting approximately 1 μL of individual liquid standard into a 1 L glass container, which was heated and the temperature was kept constant at around 90 °C. Then 1.0 mL of gas phase was injected into both gas chromatographs. All liquid standards were purchased from Fluka Chemica.

2. Results and discussion

2.1. Calibration of standards using FID

The two multicomponent standard mixtures contained individual compounds at concentrations around 1 ppm. Since this concentration is very low, both standards were precisely calibrated by using a 100 ppm n-paraffins mixture (ethane, propane, butane, pentane, hexane). Specific amounts (0.25, 0.5, 1.0, 1.5 and 2.0 mL) of the latter standard were injected into a GC-FID system respectively. Calibration curves for each individual n-paraffin were constructed (peak areas vs. injected quantities calculated in ng). Slopes represent response factors. According to the ECN theory molecular response factors are proportional with number of carbon atoms. When mass response factor is calculated out of them (division by molecular mass), approximately the same values are obtained. So on the FID 1 ng of whichever n-paraffin gives approximately the

Table 1
Data set with calculated and experimental response factors with prediction error.

Compound	Exp. RF	Calc. RF	Error (%)
1-Propanol	8.72E+03	8.84E+03	1.4
1-Butanol	1.47E+04	1.45E+04	−1.7
2-Pentanol	2.44E+04	1.65E+04	−32.4
1-Heptaldehyde	2.15E+04	2.65E+04	23.3
2-Pentanone	2.16E+04	2.03E+04	−5.9
3-Pentanone	1.91E+04	2.37E+04	24.1
4-Methyl-2-pentanone	2.25E+04	2.53E+04	12.3
2-Heptanone	2.66E+04	2.77E+04	4.2
3-Methyl-2-buten-1-ol	3.35E+04	2.54E+04	−24.2
Caprylic aldehyde	3.47E+04	2.92E+04	−16.0
Isopropyl ether	2.31E+04	2.79E+04	20.5
Hexane	3.10E+04	3.27E+04	5.3
2-Propanol	1.01E+04	1.08E+04	6.8
Diethyl ether	3.45E+04	2.16E+04	−37.4
Ethanol	4.77E+03	4.77E+03	−0.1
Acetone	1.27E+04	1.52E+04	19.3
Pentene	2.05E+04	2.43E+04	18.4
Isoprene	1.96E+04	2.27E+04	15.8
Pentane	1.87E+04	2.91E+04	55.5
cis-2-Pentene	2.28E+04	2.48E+04	8.3
2,3-Dimethylbutane	2.58E+04	3.38E+04	31.1
2-Methylpentene	2.63E+04	3.05E+04	16.1
3-Methylpentene	2.64E+04	3.04E+04	15.1
1-Hexene	2.59E+04	2.43E+04	−6.0
Hexane	3.54E+04	3.27E+04	−7.6
Methylcyclopentane	3.36E+04	3.03E+04	−9.6
2,4-Dimethylpentane	2.82E+04	3.45E+04	22.3
Cyclohexane	3.29E+04	3.56E+04	8.4
Benzene	3.99E+04	3.67E+04	−7.8
2-Methylhexane	2.53E+04	3.35E+04	32.3
2,3-Dimethylpentane	4.13E+04	3.43E+04	−17.1
3-Methylhexane	3.18E+04	3.37E+04	5.8
2,2,4-Trimethylpentane	3.67E+04	3.64E+04	−0.7
Heptane	3.32E+04	3.59E+04	8.0
Methylcyclohexane	4.18E+04	3.45E+04	−17.4
2,3,4-Trimethylpentane	4.14E+04	3.99E+04	−3.5
2-Methylheptane	3.73E+04	3.63E+04	−2.9
3-Methylheptane	4.30E+04	3.64E+04	−15.3
Octane	4.90E+04	3.87E+04	−21.0
Trichlorofluoromethane	7.71E+04	7.71E+04	0.0
1,1-Dichloroethene	4.38E+04	3.72E+04	−15.0
Methylene chloride	7.72E+04	7.95E+04	3.1
1,1-Dichloroethane	3.91E+04	3.71E+04	−5.0
cis-1,2-Dichloroethene	3.63E+04	4.01E+04	10.4
Chloroform	1.04E+05	1.01E+05	−2.9
1,1,1-Trichloroethane	4.02E+04	4.97E+04	23.8
1,2-Dichloroethane	4.23E+04	3.79E+04	−10.3
Benzene	3.88E+04	3.67E+04	−5.4
Trichloroethylene	5.54E+04	5.55E+04	0.2
Toluene	3.51E+04	3.25E+04	−7.6
1,1,2-Trichloroethane	4.19E+04	4.77E+04	13.8
1,2-Dibromoethane	7.85E+04	7.31E+04	−6.9
Tetrachloroethylene	1.06E+05	1.08E+05	1.3
Chlorobenzene	3.42E+04	3.77E+04	10.1

same response, the RF was calculated as the mean value of all slopes. With known response factor the other two standards containing hydrocarbons and chlorinated hydrocarbons were calibrated. 0.25, 0.5, 1.0, 1.5 and 2.0 mL of each multicomponent standard were injected into a GC-FID, respectively. A linear correlation (peak areas vs. injected quantities in mL) for all examined compounds was observed (correlation coefficient R^2 was in all cases greater than 0.995). By using the response factor of our FID and the ECN method [25], the mass (in ng) of each individual compound present in a multicomponent standard from its corresponding peak areas was calculated. By using the described procedure the injected mass of hydrocarbons was determined by uncertainty of up to 10%. The same procedure was applied for gathering the exact quantities (in ng) of individual hydrocarbons containing heteroatoms, however the uncertainty in the determination of injected mass was 25%.

2.2. Response factors of MSD

Specific amounts (0.25, 0.5, 1.0, 1.5 and 2.0 mL) of each multicomponent standard were injected into a GC–MS, respectively. Since both standards were previously calibrated by using the FID detector, the quantity of each compound in a mixture was known. Calibration curves for dependence of peak areas vs. injected quantities in ng were constructed. A linear correlation for all examined compounds was observed (correlation coefficient R^2 was in all cases greater than 0.995). From slopes of calibration curves response factors for each compound were obtained. The same procedure was also done for gathering the RFs of hydrocarbons containing heteroatoms. In all cases RFs represent slopes of calibration curves with the intercept value of 0. It was verified that there is no significant difference if the constant terms were considered.

2.3. Correlation of response factors

The aim of our work was to establish a quantitative correlation between chemical structure and response factor (RF) of mass selective detector. After a series of experiments and calculations linear correlation between RFs and structural descriptors was established. The CODESSA software was used to calculate structural indices and an optimal n-parameter MLR model was selected applying a stepwise selection method with up to 7 descriptors. The addition of further descriptors did not improve prediction abilities of the obtained models (at the same time the additional descriptors would increase the probability to obtain the overtrained model). The created model contained all possible structural descriptors. In Table 2 we present the best *n*-dimensional MLR model with up to seven descriptors for the modeling of response factors from chemical structure.

The best model for the prediction of response factors of MS detector from chemical structure was obtained using a 7-parameter model shown in Eq. (1).

$$RF = 6.2605e(+04) + 3.1384e(+03)a + 1.3686e(+04)b$$
$$-2.9959e(+02)c + 6.0829e(+05)d - 2.1916e(+05)e$$
$$-2.6964e(+03)f - 1.1680e(+05)g, \tag{1}$$

where *a* is the relative molecular weight, *b* average complementary information content (order 1), *c* maximum electron-electron repulsion for a C–C bond, *d* minimum net atomic charge for a F atom, *e* minimum nucleophilic reaction index for a O-atom, *f* HOMO–LUMO energy gap and *g* minimum net atomic charge for a H atom. Additional parameters did not significantly improve the quality of the prediction model. The final calculation results for all structures are presented in Fig. 1. As mentioned before capability of our model was evaluated by the leave-one-out cross-validation procedure and the RMS_{cv} error is 7000. With regard to the average experimental RF value, which is 35,700, this means that the error in prediction of response factors is just below 20%.

In Table 1 experimental vs. calculated response factors with corresponding errors for all investigated compounds are gathered. The error is expressed as the quotient of difference between calculated and experimental vs. experimental RF values.

As can be seen from Table 1 hexane and benzene appear twice. The reason for this is that both multicomponent gas standards contained benzene. In the case of the hexane, it was present in Restek multicomponent standard and it was also prepared from liquid standard. We purposely left those in our calculations so that we were able to tell if we get the same numbers in each case. From this data it is obvious that the responses of our instruments were very stable. For majority of compounds the prediction of response factors is very good, meaning that the error is in the

Table 2
The best n-parameter MLR model containing all structural descriptors with corresponding statistical parameters.

n	Parameters	r^2	q^2	s_{cv}
1	Relative molecular weight	0.660	0.617	5230
2	Relative molecular weight RNCG Relative negative charge (QMNEG/QTMINUS)	0.823	0.794	3976
3	Relative molecular weight RNCG Relative negative charge (QMNEG/QTMINUS) WNSA-1 weighted PNSA (PNSA1*TMSA/1000) [semi-MO PC]	0.860	0.829	3755
4	Relative molecular weight RNCG relative negative charge (QMNEG/QTMINUS) DPSA-1 difference in CPSAs (PPSA1-PNSA1) [semi-MO PC] ESP-Max net atomic charge for a C atom	0.880	0.841	3333
5	Relative molecular weight Average complementary information content (order 1) Max e-e repulsion for a C–C bond Min net atomic charge for a F atom Max exchange energy for a C–C bond	0.914	0.853	3055
6	Relative molecular weight Complementary information content (order 2) Max e–e repulsion for a C–C bond Min net atomic charge for a F atom Randic index (order 1) Tot point-charge comp. of the molecular dipol	0.928	0.885	2779
7	Relative molecular weight Average complementary information content (order 1) Max e–e repulsion for a C–C bond Min net atomic charge for a F atom Min nucleoph. react. index for a O-atom HOMO–LUMO energy gap Min net atomic charge for a H atom	0.939	0.904	2573

r^2, coefficient of determination; q^2, coefficient of prediction; s_{cv}, cross-validation deviation.

range of 20%. Considering the requirements for the determination of VOC in the air this is sufficient. The worst predicted compounds with the error above 30% are 2-pentanol, diethyl ether, pentane, 2,3-dimethylbutane and 2-methylhexane. One of the reasons for inaccurate prediction of the last three mentioned VOCs could be because few of the peaks in chromatograms of multicomponent standards were overlapping a bit despite their separation. As the result of this the area measurement was not entirely precise and small inaccuracies could further intensify during our calculation procedure and yield worse results. Our prediction model shows good results for hydrocarbons containing heteroatoms. This group of compounds is a more complex system in comparison to regular hydrocarbons, since they are more polar. Considering this the error in the prediction of 2-pentanol and diethyl ether between 30 and 40% is somewhat understandable. Despite this our prediction

model with selected calculated descriptors still managed to represent the whole system with hydrocarbons containing heteroatoms contentedly.

In order to test the applicability of our prediction model during longer period of time several experiments were performed. The base signal of MSD changes during time, meaning that the response factors for the same compounds will not always be the same. The idea of our work is to calibrate the instrument before usage to just few compounds (which are present in the prediction model) and then shift the whole calibration curve according to the measured responses. In order for this to work the trend of response factor's changes must be the same for all compounds. Ten different VOCs (2,3-dimethylbutane, 3-methylpentene, hexane, benzene, 3-methylhexane, 2,2,4-trimethylpentane, heptane, methylcyclohexane, 2,3,4-trimethylpentane and 3-methylheptane) were injected
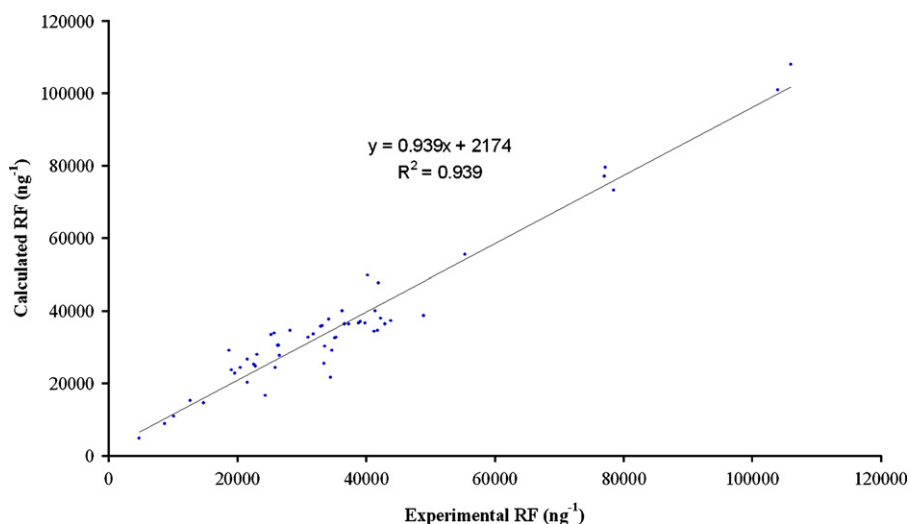

Fig. 1. Calculated vs. experimental RF.

Table 3
Average coefficients for all possible combinations of 10 VOCs with according standard deviations ($n = 8$).

Average ± stdev	A[a]	B[b]	C[c]	D[d]	E[e]	F[f]	G[g]	H[h]	I[i]	J[j]
A[a]	1	0.95 ± 0.04	0.99 ± 0.07	1.03 ± 0.04	1.24 ± 0.06	1.59 ± 0.08	1.17 ± 0.06	1.47 ± 0.10	1.55 ± 0.11	1.53 ± 0.12
B[b]	1.06 ± 0.05	1	1.05 ± 0.06	1.09 ± 0.03	1.31 ± 0.06	1.68 ± 0.08	1.24 ± 0.07	1.56 ± 0.10	1.64 ± 0.12	1.62 ± 0.13
C[c]	1.01 ± 0.07	0.96 ± 0.05	1	1.04 ± 0.07	1.26 ± 0.10	1.61 ± 0.09	1.18 ± 0.08	1.49 ± 0.14	1.57 ± 0.13	1.55 ± 0.16
D[d]	0.98 ± 0.04	0.92 ± 0.03	0.97 ± 0.06	1	1.21 ± 0.05	1.55 ± 0.08	1.13 ± 0.06	1.43 ± 0.09	1.51 ± 0.11	1.49 ± 0.13
E[e]	0.81 ± 0.04	0.76 ± 0.03	0.80 ± 0.06	0.83 ± 0.04	1	1.28 ± 0.06	0.94 ± 0.04	1.18 ± 0.03	1.25 ± 0.08	1.23 ± 0.09
F[f]	0.63 ± 0.03	0.60 ± 0.03	0.62 ± 0.04	0.65 ± 0.04	0.78 ± 0.04	1	0.74 ± 0.03	0.93 ± 0.05	0.97 ± 0.03	0.96 ± 0.06
G[g]	0.86 ± 0.04	0.81 ± 0.05	0.85 ± 0.05	0.88 ± 0.05	1.06 ± 0.04	1.36 ± 0.06	1	1.26 ± 0.05	1.33 ± 0.08	1.31 ± 0.10
H[h]	0.68 ± 0.04	0.65 ± 0.04	0.68 ± 0.06	0.70 ± 0.04	0.85 ± 0.02	1.08 ± 0.06	0.80 ± 0.03	1	1.05 ± 0.07	1.04 ± 0.08
I[i]	0.65 ± 0.04	0.61 ± 0.04	0.64 ± 0.05	0.67 ± 0.05	0.80 ± 0.05	1.03 ± 0.04	0.76 ± 0.05	0.95 ± 0.06	1	0.99 ± 0.04
J[j]	0.66 ± 0.06	0.62 ± 0.05	0.65 ± 0.07	0.68 ± 0.06	0.82 ± 0.06	1.04 ± 0.06	0.77 ± 0.07	0.97 ± 0.07	1.01 ± 0.05	1

[a] A, 2,3-dimethylbutane.
[b] B, 3-methylpentene.
[c] C, hexane.
[d] D, benzene.
[e] E, 3-methylhexane.
[f] F, 2,2,4-trimethylpentane.
[g] G, heptane.
[h] H, methylcyclohexane.
[i] I, 2,3,4-trimethylpentane.
[j] J, 3-methylheptane.

daily over a period of two weeks and ratios between all combinations of those compounds were calculated. In other words response of every individual VOC was divided with all other VOCs in order to minimize the error in measured response of every individual compound and as a result 100 coefficients were calculated. In Table 3 results are presented as average values of 8 separate measurements and their corresponding standard deviations. The average error between all coefficients, expressed as a relative standard deviation, is just 6%. From those results it is evident that the instrument only needs to be calibrated to few compounds and then accordingly shifted prediction model can be applied for the prediction of response factors.

When using our developed model for the prediction of VOC's response factors of individual MS detector, several steps must be taken into consideration. Firstly one must check the stability of MS detector and then RFs of few VOCs, that are present in the prediction model, must be determined. If the ratios between experimentally determined and VOCs in the model are the same, then the model can be used (if needed the whole model can be shifted according to the ratio between our RFs and newly determined RFs). To quantify individual compound, seven structural descriptors, that are used in developed model, must be calculated. Then from Eq. (1) RF can be obtained. If the ratios between experimentally determined and VOCs in the model are not matching, then the whole experimental procedure must be done. In other words experimental RFs for all used VOCs must be obtained and then the QSPR correlation between chemical structures and RFs must be established as stated in "creation of the models" part. Then the newly developed model can be used as already described.

3. Conclusions

In this paper our study of calibrating of GC/MS instrument is presented. An accurate quantification represents one of the key problems in analysis which is even more pronounced when dealing with non-target analysis. A quantitative structure–property relationship (QSPR) model for the prediction of MS response factors was developed for the quantification of volatile organic compounds in the atmosphere. Previously there was no numerical model, which would enable prediction of MS response factors from chemical structures of compounds. The CODESSA software was used to calculate structural indices and for selection of an optimal n-parameter MLR model. We established a linear correlation between RFs and structural descriptors. The best results for the prediction of RFs from

chemical structure were obtained by using a 7-parameter model. The average capability of our model was evaluated by the leave-one-out cross-validation procedure and was just below 20%, which is sufficient for the determination of VOCs in the air, where concentrations of various organic compounds are found in ppb or even in ppt levels.

Current evaluation of the proposed procedure does not take into account serious matrix problems in cases when liquid samples are analyzed. The proposed estimation procedure is insensitive for the changes of sensitivity of the MS detector as long as sensitivity is influenced the same way along the whole m/z axis. As long as this is true, one can use the same model for different chromatographic procedures. Such small matrix effects can be expected in cases of atmospheric research, where relatively clean samples (VOCs in ppb to ppt ranges) are analyzed. This is also the application for which the procedure was developed.

The solvent can affect the whole chromatographic line from liner to the ion source and therefore change the sensitivity of the MS detector. It is known that such effects will usually influence just part of the m/z scale. The MS instrument will become insensitive just for lower or higher masses. The proposed procedure cannot compensate for such effect. In such instances the model cannot be transferred from one analytical procedure to another.

At the end it should be noted that the quantification of the organic compounds using described procedure is possible only when the structures of the unknowns have been clarified using for instance electron impact MS spectra library search or comparison with the retention indices database. The proposed procedure is not convenient if exact quantification of specific compounds is required since it would probably take the same time to calibrate the instrument with that particular compound separately. However in cases of non target analysis where some new compounds are identified during the reprocessing step, the described procedure enables you to tentatively quantify the compounds identified by the MS library search. In such cases the original sample is usually no longer available so additional calibration of the instrument is not possible anymore. This is especially true in atmospheric chemistry research when analyst is faced with the challenge to identify and quantify as many as possible organic compounds among few hundred chromatographic peaks.

The further study will be needed to simplify the modeling step before the modeling procedure will be applicable for the routine laboratories. The study should address simplification in cases of

high structural similarities or just simplified structural representation of chemical compounds.

Acknowledgments

The authors acknowledge the financial support by the Ministry of Higher Education, Science and Technology of the Republic of Slovenia (Grant P1-0153) and the Ministry of Science, Education and Sports of the Republic of Croatia (Grant 098-0982915-2947) and to both Ministries for granting bilateral project (BI-HR/09-10-38).

References

[1] B.J. Finlayson-Pitts, J.N. Pitts Jr., Chemistry of the Upper and Lower Atmosphere, Academic Press, London, 1999.
[2] A.J. Haagen-Smit, M.M. Fox, J. Air Pollut. Control Assoc. 4 (1954) 105.
[3] Y. Choi, S.H. Ehrman, Atmos. Environ. 38 (2004) 775.
[4] M. Ulman, Z. Chilmonczyk, Chem. Anal. (Warsaw) 52 (2007) 173.
[5] S. Saito, I. Nagao, H. Kanzawa, Atmos. Environ. 43 (2009) 4384.
[6] K. Na, K. Moon, Y.P. Kim, Atmos. Environ. 39 (2005) 5517.
[7] H. Guo, S.C. Lee, L.Y. Chan, W.M. Li, Environ. Res. 94 (2004) 57.
[8] R.E. Clement, P.W. Yang, Anal. Chem. 69 (1997) 251R.
[9] B. Clemens, R.G. Lewis, Sampling for Organic Compounds, Principles of Environmental Sampling, American Chemical Society, United States of America, 1988, 287 pp.
[10] L. Fox, Anal. Chem. 69 (1997) 1R.
[11] S.A. Batterman, G.Z. Zhang, M. Baumann, Atmos. Environ. 32 (1998) 1647.
[12] M.H. Hiatt, Anal. Chem. 67 (1995) 4044.
[13] M.H. Hiatt, D.R. Youngman, J.R. Donnelly, Anal. Chem. 66 (1994) 905.
[14] M.H. Hiatt, C.M. Farr, Anal. Chem. 67 (1995) 426.
[15] Th. Knobloch, A. Asperger, W. Engewald, Fresenius J. Anal. Chem. 359 (1997) 189.
[16] D. Waterman, B. Horsfield, F. Leistner, K. Hall, S. Smith, Anal. Chem. 72 (2000) 3563.
[17] E. Massold, C. Bähr, T. Salthammer, S.K. Brown, Chromatographia 62 (2005) 75.
[18] S.N. Sinha, P.K. Kulkarni, N.M. Desai, S.H. Shah, G.M. Patel, M.M. Mansuri, D.J. Parikh, H.N. Saiyed, J. Chromatogr. A 1065 (2005) 315.
[19] Ö.O. Kuntasal, D. Karman, D. Wang, S.G. Tuncel, G. Tuncel, J. Chromatogr. A 1099 (2005) 43.
[20] A.R. Katritzky, E.S. Ignatchenko, R.A. Barcock, V.S. Lobanov, Anal. Chem. 66 (1994) 1799.
[21] A.R. Katritzky, V.S. Lobanov, M. Karelson, Chem. Soc. Rev. 24 (1995) 279.
[22] P. Liu, W. Long, Int. J. Mol. Sci. 10 (2009) 1978.
[23] C. Nantasenamat, C. Isarankura-Na-Ayudhya, T. Naenna, V. Prachayasittikul, EXCLI J. 8 (2009) 74.
[24] J.C. Sternberg, W.S. Gallaway, D.T.L. Jones, Gas Chromatography, Academic Press, New York, 1962 (Chapter 18).
[25] J.T. Scanlon, D.E. Willis, J. Chrom. Sci. 23 (1985) 333.
[26] H.Y. Tong, F.W. Karasek, Anal. Chem. 56 (1984) 2124.
[27] A.D. Jorgensen, K.C. Picel, V.C. Stamoudis, Anal. Chem. 62 (1990) 683.
[28] H. Yieru, O. Qingyu, Y. Weile, Anal. Chem. 62 (1990) 2063.
[29] M. Jalali-Heravi, M.H. Fatemi, J. Chromatogr. A 825 (1998) 161.
[30] M. Jalali-Heravi, M.H. Fatemi, J. Chromatogr. A 897 (2000) 227.
[31] M. Jalali-Heravi, E. Noroozian, M. Mousavi, J. Chromatogr. A 1023 (2004) 247.
[32] G. Arh, L. Klasinc, M. Veber, M. Pompe, Acta Chim. Slov. 57 (2010) 581.
[33] J.J.P. Stewart, J. Comput. Aided Mol. Des. 4 (1990) 1.
[34] A.R. Katritzky, V.S. Lobanov, M. Karelson, CODESSA Training Manual, University of Florida, Gainesville, Florida, 1995.
[35] A.R. Katritzky, L. Mu, V.S. Lobanov, M. Karelson, J. Phys. Chem. 100 (1996) 10400.
[36] A.R. Katritzky, E.V. Gordeeva, J. Chem. Inf. Comput. Sci. 33 (1993) 835.
[37] M. Karelson, V.S. Lobanov, A.R. Katritzky, Chem. Rev. 96 (1996) 1027.
[38] M. Veber, M. Pompe, H. Baša-Češnik, J. Inst. Sci. Technol. Balikesir Univ. 3 (2001) 125.