

**APPLICATION OF CHEMOMETRICS METHODS FOR THE EVALUATION  
OF DATA FROM THE STUDY OF LONG TERM STABILITY OF THE  
DETECTOR RESPONSE FOR TEST SOLUTIONS USED FOR CALIBRATION  
IN ION CHROMATOGRAPHY <sup>†</sup>**

**Janja Turšič, Marjana Novič, and Milko Novič**

*National Institute of Chemistry, Ljubljana, Hajdrihova 19, Slovenia*

<sup>†</sup>This paper is based on a presentation at The Fourth Slovenian Symposium on Separation Techniques meeting held in Novo Mesto, Slovenia during October 3-4<sup>th</sup>, 2002.

*Received 16-12-2002*

### **Abstract**

A set of experiments performed to study long-term stability of the working standard solutions in ion chromatography (IC) was examined using different statistical methods. Random variations in the detector response using sufficiently stable and homogeneous samples were observed. Six anions were included in the study: Cl<sup>-</sup>, NO<sub>2</sub><sup>-</sup>, Br<sup>-</sup>, NO<sub>3</sub><sup>-</sup>, HPO<sub>4</sub><sup>2-</sup> and SO<sub>4</sub><sup>2-</sup> in concentration ranges 2 mgL<sup>-1</sup> (chloride, nitrite), 5 mgL<sup>-1</sup> (bromide, nitrate) and 10 mgL<sup>-1</sup> (phosphate and sulphate). Two duplicates were prepared from stock solution by dilution in regular time intervals of one week. One of the duplicates was stored in the refrigerator, while the other was kept in dark at room temperature. After separation, both conductivity and spectrophotometric detection (UV/Vis) were applied. The influence of analyst was minimized using completely automated and thermostated system. The duration of performed test was 3 months. It was shown with pair-wise comparisons of group means using t-test that there are significant fluctuations in the detector response in the whole time period of 3 months, no matter whether the measurements were performed on aged samples or on fresh samples prepared each week from the stock solution. The application of the ANOVA test also showed that groups of experiments performed on equally aged samples in some cases differ significantly. The frequent variation of the recorded detector signals does not origin mainly from aging of samples.

### **Introduction**

The calibration of an instrument is an essential part of every analytical method. For the calibration we use either a reference material of known composition, or calibrant, in which the property of interest is well characterized. Therefore the reagents, reference material and calibrants have to be under control regarding their chemical and physical decomposition. Additionally, different sources of errors, such as the analyst's manipulation with samples during dilution process, laboratory conditions, instrument's

hardware instability etc., introduce a random variation in performance of the analytical method using sufficiently stable and homogeneous samples.

The aim of our study was to evaluate the uncertainty in IC by monitoring the detector response obtained for different anions (peak area) in test solutions. They were prepared (in regular time intervals) from stock solution by dilution and contained common anions ( $\text{Cl}^-$ ,  $\text{NO}_2^-$ ,  $\text{Br}^-$ ,  $\text{NO}_3^-$ ,  $\text{HPO}_4^{2-}$  and  $\text{SO}_4^{2-}$ ). One half of samples were stored in the refrigerator and the other half was kept in dark at room temperature. After separation on column, simultaneous conductivity and spectrophotometric (UV/Vis) detections were applied. The duration of performed test was 3 months. The influence of analyst was minimized using completely automated and thermostated system.

Statistical tests were applied in order to study the origin of fluctuations in the detector response. With the ANOVA<sup>1</sup> test (ANalysis Of VAriance) the comparison between the groups of experiments performed on equally aged samples was made. With pair-wise comparisons of group means using more sensitive t-test it was investigated, whether there were significant fluctuations in the detector response in the whole time period of 3 months. The test was made between the smallest groups, each containing 6 or 5 replicates, and also on larger groups, made on the basis of equal age of samples or on the basis of common sample preparation or common week of measuring.

### **Experimental**

The used experimental set-up was a Dionex DX500 chromatography system (Sunnyvale, CA, USA), consisting of a GP50 quaternary gradient pump, an EG40 eluent generator module, a LC25 thermostated chromatography module and an ED40 conductivity detector module. As an additional external detector an UV/Vis spectrophotometer (Spectras-Physics SpectraSYSTEM, Fremont, USA) working at 210 nm was used. The separation columns and suppressor used included a Dionex IonPac AG4A-SC (4x50 mm) guard column, an IonPac AS4A-SC (4x250 mm) separation column, and an ASRS-ULTRA-4mm suppressor working in recycle mode. The sample loop volume was 50  $\mu\text{L}$ . The eluent flow rates were 2 mL/min. The entire chromatography system was controlled using Dionex PeakNet 5.1 chromatography software. UV/Vis detector responses were collected by PeakNet 5.1 through UI20 (Dionex, Sunnyvale, USA). Sample introduction was automated using a 6-port selection

valve (Rheodyne, CA, USA) and a peristaltic pump Minipuls3 (Gilson, France), both controlled by PeakNet 5.1 through I/O installed on GP50 gradient pump.

All the reagents used in this study were of p.a. grade (Merck, Darmstadt, Germany). The stock solution was prepared from corresponding sodium and potassium salts and contained  $100 \text{ mgL}^{-1} \text{ Cl}^-$  and  $\text{NO}_2^-$ ,  $250 \text{ mgL}^{-1} \text{ Br}^-$  and  $\text{NO}_3^-$ , and  $500 \text{ mgL}^{-1} \text{ HPO}_4^{2-}$  and  $\text{SO}_4^{2-}$ . The test solutions were prepared from stock solution by dilution (1:50) in the concentration ranges  $2 \text{ mgL}^{-1}$  (chloride, nitrite),  $5 \text{ mgL}^{-1}$  (bromide, nitrate) and  $10 \text{ mgL}^{-1}$  (phosphate and sulphate) with Milli-Q water ( $18 \text{ MWcm}^{-1}$ , Millipore, Bedford, MA, USA). Stock solution as well as test solutions was filtered through  $0.2 \mu\text{m}$  filter (Millipore, Bedford, MA, USA) in order to prevent potential biodegradation. As the eluent carbonate/bicarbonate buffer ( $1.8 \text{ mM Na}_2\text{CO}_3 / 1.7 \text{ mM NaHCO}_3$ ) was used.

### Results and discussion

The long-term stability test of the calibration in IC was performed in the time-period of three months. The samples were prepared in one-week time intervals from stock solution. Between two duplicates, one was stored in refrigerator and the other one was kept in dark at room temperature. For the first five weeks, each sample was analysed in 6 replicates every week, which produced 12 measurements for each anion in the first week, 24 in the second, 36 in the third, and so on until the sixth week, in which the number of replicates was lowered to 5 because the number of samples was already extended to 12. In the last week there were 80 measurements for each anion. As can be seen from Table 1, in few cases less than 5 replicates were recorded, which was due to accidental troubles with the instrument during recording. Altogether 510 measurements were performed. For three anions (nitrite, bromide and nitrate) besides conductivity, data obtained using UV/Vis detection were also available. In Table 1 number of replicates for measurement of samples prepared and recorded in different time-period is given. It can be seen from Table 1 that fresh samples were being prepared for 8 weeks (first row), while the measurements were going on for 14 weeks (first column). In columns 2 to 9 there are replicates of the sample prepared in the week indicated in the first row, and consequently analysed during the whole period of the experiment (e.g. the second column contains replicates of the sample prepared in the first week and measured for 14 weeks). The age of each measured sample can be calculated by subtracting the

*week\_preparation* from the *week\_measurement* (numbers from the first row and first column in Table 1, respectively) If these two numbers are equal, the age is 0 weeks, this is the fresh sample prepared in the corresponding week.

To study the dependence of peak area of the detector response on the age of sample, the samples of the same age (constant difference of the *week\_measurement-week\_preparation*, elements parallel to the diagonal in Table 1) were grouped. Consequently, the first group (freshly prepared samples) contained two subgroups of 45 measurements, samples kept in the refrigerator or at room temperature, respectively, while the 13<sup>th</sup> group contained two subgroups of 10 measurements. In Figures 1 and 2 it is shown how the average peak area and relative standard deviation *rsd* (standard deviation *s*, expressed in percentages as a fraction of the mean value, see Eq. 1) for nitrite and bromide ions vary with the sample aging. The results are given for UV/Vis detector.

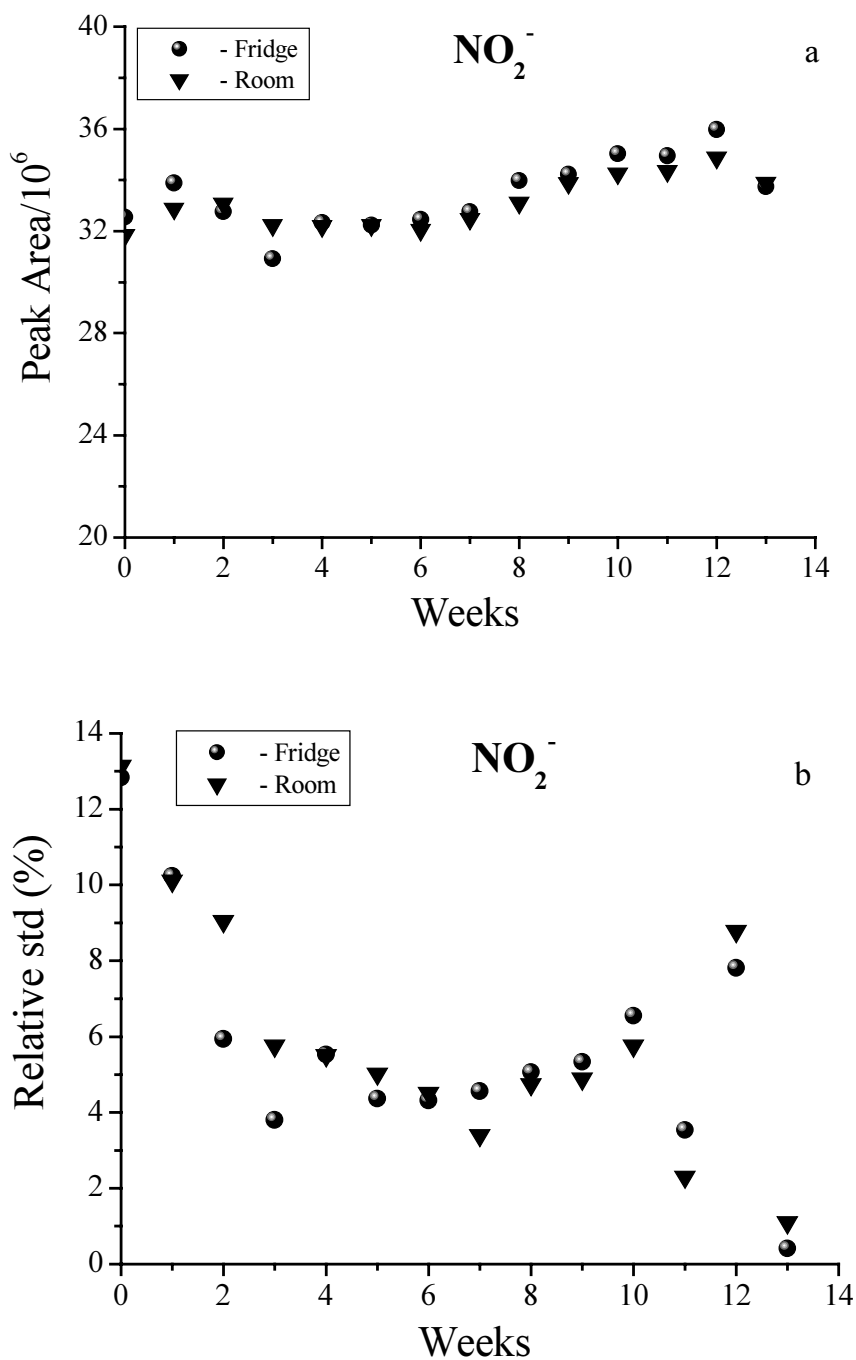
**Table 1.** Time-schedule for sample preparation and measurements (replicates).

Week - sample preparation Week - measurement	1	2	3	4	5	6	7	8
1	6/6*							
2	6/6	6/6						
3	6/6	6/6	6/6					
4	6/6	6/6	6/6	6/6				
5	6/6	6/6	6/6	6/6	6/6			
6	5/5	5/5	5/5	5/5	5/5	5/5		
7	5/5	5/5	5/5	5/5	5/5	5/5	5/5	
8	5/5	5/5	5/5	5/5	5/5	5/5	5/5	5/5
9	5/5	5/5	5/5	5/5	5/5	5/5	5/5	5/5
10	5/5	3/5	5/5	5/5	5/5	5/5	5/5	5/5
11	5/2	1/5	5/5	5/5	5/5	5/5	5/5	5/5
12	5/5	5/5	5/5	5/5	5/5	4/4	5/5	5/5
13	5/5	5/5	5/5	5/5	5/5	5/5	5/5	5/5
14	5/5	5/5	5/5	5/5	5/5	5/5	5/5	5/5

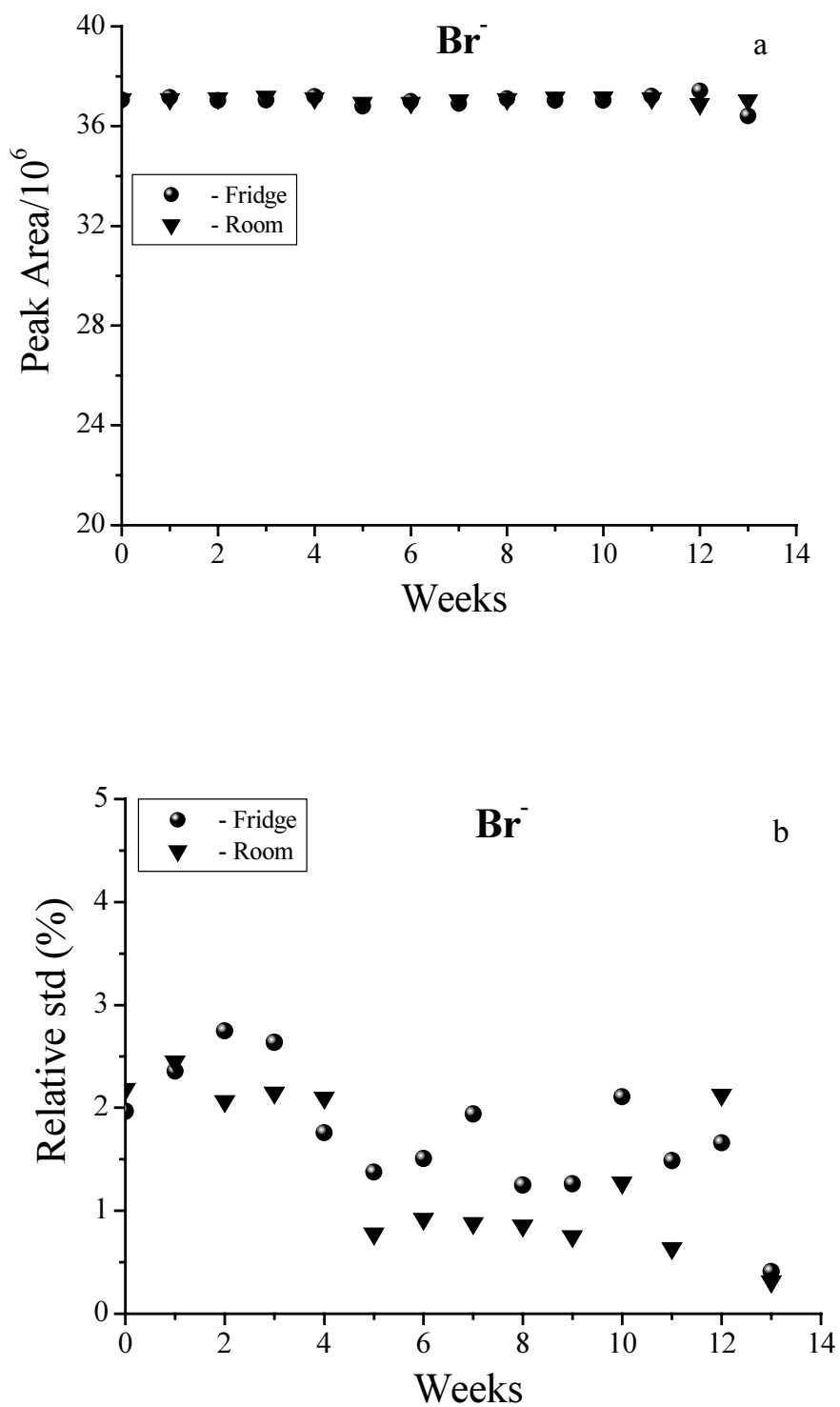
\* 6/6 stays for 6 replicates from the flask stored in the refrigerator (left) and 6 replicates from the flask stored at room temperature (right).

$$rsd = \frac{100 \times s}{\bar{x}} = \frac{100}{\bar{x}} \times \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{Eq. 1}$$

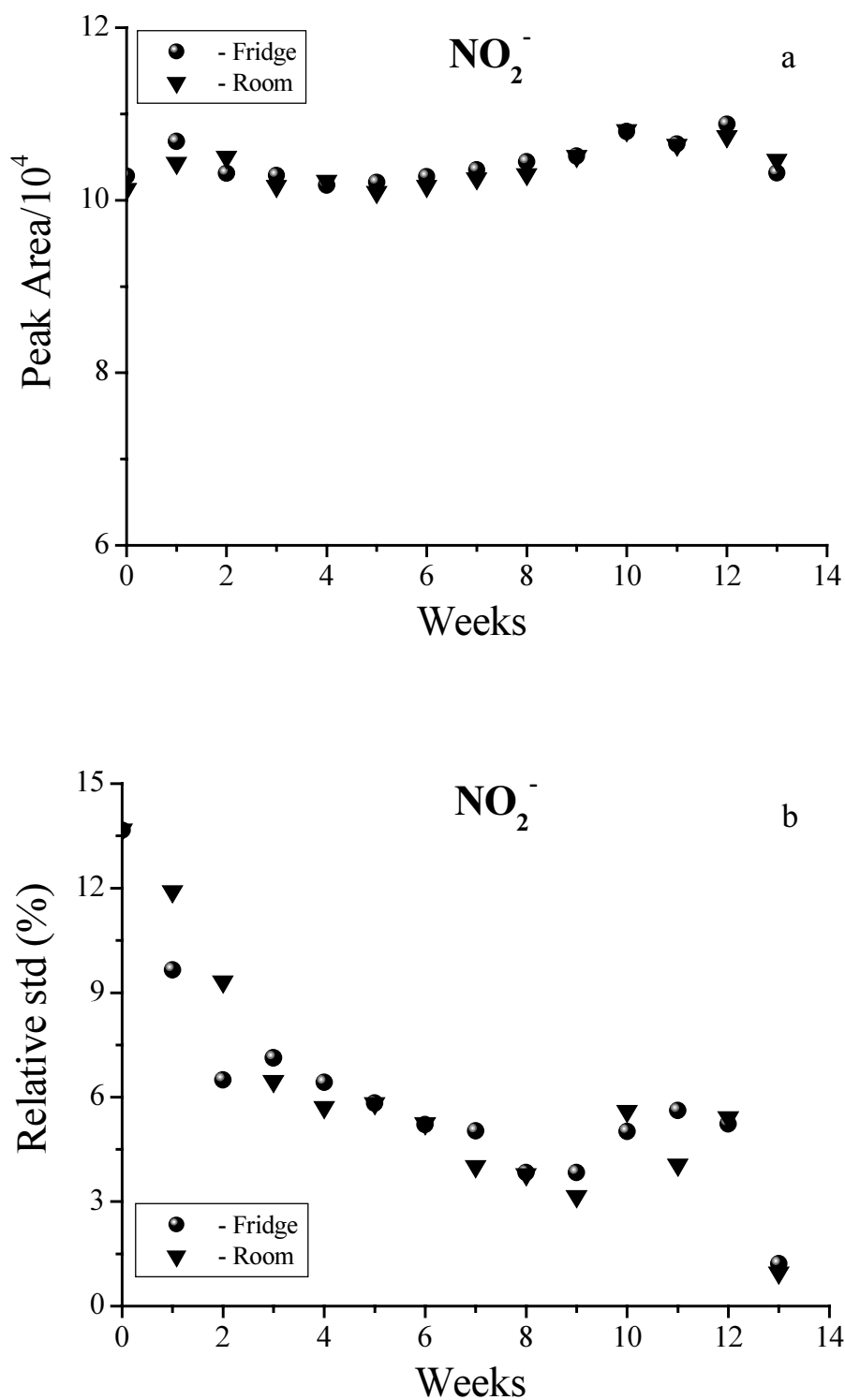
The same procedure of merging equally old samples was done also for the case of conductivity detector. Corresponding plots are shown in Figures 3 and 4.



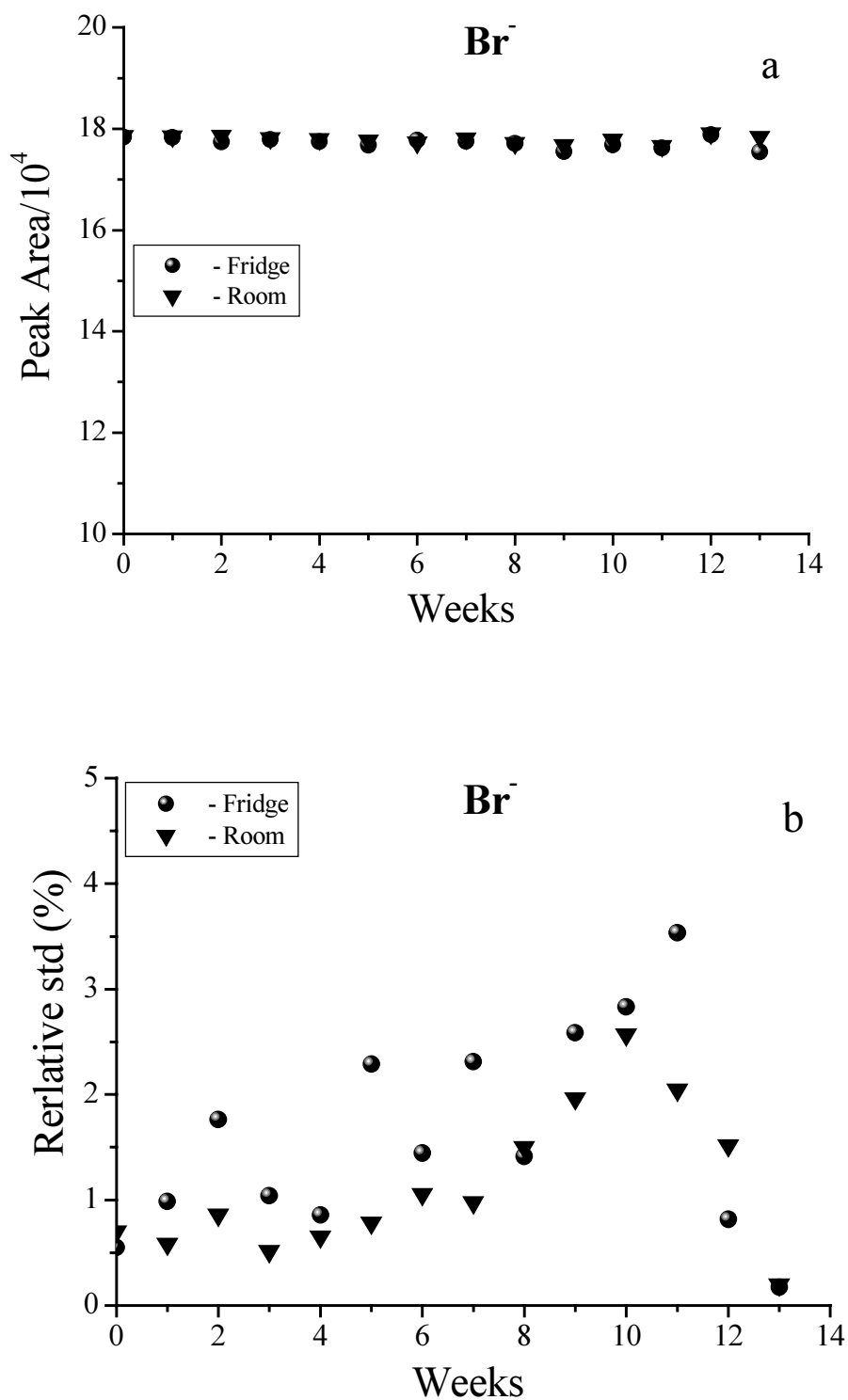
**Figure 1.** Average peak areas (a) and relative standard deviations (b) for  $\text{NO}_2^-$  ions, UV/Vis detector. The circles and triangle data markers stand for samples kept in the refrigerator or at room temperature, respectively.



**Figure 2.** Average peak areas (a) and relative standard deviations (b) for  $\text{Br}^-$  ions, UV/Vis detector. Other data are the same as marked in Fig. 1.



**Figure 3.** Average peak areas (a) and relative standard deviations (b) for  $\text{NO}_2^-$  ions, conductivity detector. Other data are the same as marked in Fig. 1.



**Figure 4.** Average peak areas (a) and relative standard deviations (b) for  $\text{Br}^-$  ions, conductivity detector. Other data are the same as marked in Fig. 1.



From Figs. 1 and 3 and from Figs. 2 and 4 it can be seen that UV/Vis and conductivity detector produced similar results. In both cases the relative standard deviations are around 6% for nitrite and below 3% for bromide. Large relative standard deviation was also observed for nitrate, while the relative standard deviations for all other ions were around 2%.

A comparison of groups of measurements (given in Table 1, each group is one cell) with pair-wise t-test demonstrated significant differences between them in most of the checked pairs for the nitrite and nitrate ions, while for the other ions the pairs of groups which did not differ significantly were more frequent. It was also observed that the number of significantly different pairs did not coincide with the time period difference between the two samples. The pairs of groups showing significant difference were randomly distributed, if the position of pairs were checked in the Table 1. With other words, the differences of the peak areas of the detector signals do not correlate with the aging of samples. Additionally, most of the t-tests performed between the groups of samples prepared and measured at the same day (pairs on the diagonal of Table 1) were significant, which demonstrates that fluctuations originating from the dilution of stock solution are higher than those of the detector responses. It is known that pair-wise t-tests do not show a general picture about the integral mean – we can only judge about pairs of samples, while the comparison of only one sample sequentially with all others is biased to this first sample. In order to investigate the assumption about the origin of fluctuations of the detector response, we used the ANOVA in which a non-biased comparison of multiple means can be done. The ANOVA test was performed for the groups of samples of the same age. The 14 weeks period of testing the stability of test solutions for calibration in ion chromatography generated 14 groups of measurements of equally old samples, kept in the refrigerator. Additional 14 groups were obtained from the samples kept at room temperature. In the first group there are 45 measurements of fresh samples (prepared and measured in the same week), in the second group 44 measurements of one week old samples, and so on to the last, 14<sup>th</sup> group, in which there were only 5 measurements – the replicates of one, 13 weeks old sample. The number of measurements in each group can be controlled by counting of replicates written in Table 1, in the cells parallel to the diagonal elements. It means that the cells of Table 1 containing fresh, i.e. zero weeks old samples, are in the  $i$ -th column,  $j$ -th row, if

$i=j$ ; for one week old samples the relation between the column and row number should be  $i=j+1$ , for two weeks old samples  $i=j+2$ , and so on to the last, 14<sup>th</sup> group for 13 week old sample, which contains only 5 measurements given in the cell  $i=1, j=14$  ( $i=j+13$ ). The results of the ANOVA test between these 14 groups, containing 45, 44, 43, 42, 41, 40, 39, 35, 28, 21, 20, 15, 10, and 5 measurements are gathered in Table 2.

**Table 2.** F-values from the ANOVA test for the groups of samples of the same age.

Place of storage	F-values calculated for different anion measurements $F_{13,414}^{0.05} = 1.88^*$					
	Cl <sup>-</sup>	NO <sub>2</sub> <sup>-</sup>	Br <sup>-</sup>	NO <sub>3</sub> <sup>-</sup>	HPO <sub>4</sub> <sup>2-</sup>	SO <sub>4</sub> <sup>2-</sup>
Conductivity detector						
Fridge	4.76	2.07	2.06	1.06	1.81	1.93
Room T	7.21	1.96	3.25	1.94	2.02	1.65
UV/Vis detector						
Fridge		5.12	1.21	1.77		
Room T		3.53	0.58	3.01		

\* Theoretical F-value at significance level  $\alpha=0.05$ , degrees of freedom  $df_b=13$ ,  $df_w=414$ .

The theoretical F-value for 14 groups and altogether 428 measurements, at 0.05 significance level is 1.88. From Table 2 it can be seen that the calculated  $F$  values showing the ratio of variances between the groups of equally old samples and within all groups are not always below the theoretical value at 0.05 significance level. For certain anions, the ANOVA-test is highly significant (Cl<sup>-</sup>, NO<sub>2</sub><sup>-</sup>, Br<sup>-</sup> for conductivity detector), what means that the differences between the 14 groups are significant, one or more groups of equally old samples show a bias.

As already stressed, the variance in the detector response could be ascribed not only to the aging of samples (VAR-age), but also to the manipulation with the solution during the sample preparation (dilution of the stock solution – VAR-sample), and to the instability of the instrument (VAR-instrument). The former variance (due to dilution) is excluded from the group of measurements of the same sample (columns in Table 1), while the latter one (due to instrument instability) is minimized in the group of measurements performed in one day - the same week of measurement (rows in Table 1).

In order to compare the three sources of the variance, the ANOVA-test was repeated on the same 428 measurements, but the groups were made according to the sample (8 groups, columns in Table 1) or according to the week of measurement (14 groups, rows in Table 1). The  $F$  value was calculated only for  $\text{NO}_2^-$  ion using responses of UV/Vis detector for samples stored in the refrigerator.  $F_{VAR-age}=5.12$  can be found in the Table 2, while the two additional ANOVA tests gave  $F_{VAR-sample}=23.78$  and  $F_{VAR-instrument}=12.55$ .

### Conclusions

The study of long-term stability of working standard solutions in ion chromatography was performed to control the stability of test solutions for calibration. The stability of the detector response in IC during the three months period was also tested. After performing all the measurements, a huge amount of data was collected. Frequent variations of signals were observed. Statistical tests (t-test, F-test, ANOVA) were performed to find out whether the variations are correlated with the aging of test samples. From pair-wise t-test it was concluded that there are significant variations between the means of replicates in all cases, also when comparing the groups of samples of the same age. Relative standard deviation within the replicates was in principle significantly lower than within any other grouping of measurements. It was also shown that the determination of nitrite and nitrate was less precise (relative s over 5%) than of the other four ions. The aging of samples introduced less variance to the determination of the studied ions than the sample preparation or the instability of the instrument.

The obtained results demonstrated fluctuations in the detector responses, which could not be ascribed to chemical changes in sample composition. However, it was not straightforward to see if there were any trends in the observed fluctuations without applying some statistical tests, because the number of the measured samples increased to 90 for each ion in the last week. With pair-wise comparisons of group means using t-test it was demonstrated that there are significant fluctuations in the detector response in the whole time period of 3 months. The results of the applied ANOVA test showed that groups of experiments performed on equally aged samples do not differ significantly in 4 cases of the performed test. In 8 comparisons the calculated  $F$  values are of about the same magnitude as the theoretical one, while in the remaining 6 cases they are

considerably higher. The variations in the detector response that origin from other sources than sample instability demand rather frequent re-calibration of the instrument in order to retain the uncertainty of the applied IC system low, especially for the nitrite and nitrate.

### Acknowledgements

The Slovenian Ministry of Education, Science and Sports, for which it is kindly acknowledged, supported this work.

### References and Notes

1. D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. S. Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier, Amsterdam, 1997, pp 59–74.

### Povzetek

Preko daljšega časovnega obdobja smo spremljali stabilnost delovnih raztopin, ki se uporabljajo pri ionski kromatografiji. Rezultate smo ovrednotili z različnimi statističnimi testi. Ugotovili smo, da pri meritvah homogenih in dovolj stabilnih vzorcev prihaja do naključnih variacij detektorskega odziva. V študijo je bilo vključenih šest anionov:  $\text{Cl}^-$ ,  $\text{NO}_2^-$ ,  $\text{Br}^-$ ,  $\text{NO}_3^-$ ,  $\text{HPO}_4^{2-}$  in  $\text{SO}_4^{2-}$  v koncentracijskem območju  $2 \text{ mgL}^{-1}$  (klorid, nitrit)  $5 \text{ mgL}^{-1}$  (bromid, nitrat) in  $10 \text{ mgL}^{-1}$  (fosfat in sulfat). Iz osnovne raztopine smo v tedenskih časovnih intervalih z usreznim redčenjem pripravili dva testna vzorca. En vzorec smo hranili v hladilniku, drugega pa v temi pri sobni temperaturi. Separaciji na analizni koloni je sledila konduktometrična in spektrofotometrična (UV/Vis) detekcija. Vpliv analitika na potek meritev je bil zaradi uporabe popolnoma avtomatiziranega in termostatiranega sistema minimalen. Stabilnost detektorskega signala za posamezne anione v testnih raztopinah smo spremljali tri mesece. Primerjave povprečnih detektorskih odzivov med različnimi skupinami z uporabo t-testa so pokazale, da v celotnem obdobju trajanja testa prihaja do signifikantnih nihanj, ne glede na to ali med sabo primerjamo odzive sveže pripravljenih ali že postaranih testnih raztopin. Tudi rezultati ANOVE so pokazali, da so med enako starimi vzorci v primeru nekaterih anionov signifikantne razlike. Na osnovi te študije bi lahko zaključili, da nihanje dektorskega signala ne izhaja samo iz staranja oz. nestabilnosti testnih vzorcev.