

Towards adaptive method for peak migration time correction: discretization period in electropherograms

Tomas Drevinskas¹,

Audrius Maruška^{1*},

Gintarė Naujokaitytė¹,

Laimutis Telksnys^{2,3},

Mihkel Kaljurand^{4,1},

Vidmantas Stanys⁵,

John Cowles⁶,

Jelena Gorbatova⁴

¹ *Instrumental Analysis Open Access Centre,
Faculty of Natural Sciences,
Vytautas Magnus University,
8 Vileikos Street,
44404 Kaunas, Lithuania*

² *Department of Systems' Analysis,
Faculty of Informatics,
Vytautas Magnus University,
8 Vileikos Street,
44404 Kaunas, Lithuania*

³ *Institute of Data Science
and Digital Technologies,
Vilnius University,
12 Goštauto Street,
01108 Vilnius, Lithuania*

⁴ *Department of Chemistry,
Faculty of Sciences,
Tallinn University of Technology,
Akadeemia tee 15,
12618 Tallinn, Estonia*

⁵ *Institute of Horticulture of Lithuanian Research Center
for Agriculture and Forestry,
54333 Babtai,
Kaunas Distr., Lithuania*

⁶ *Département Biologique,
Institut Universitaire de Technologie,
University of Angers, 4 bd Lavoisier,
49016 Angers, France*

Capillary electrophoresis often causes unrepeatability of peak migration times in the electropherogram due to changes of electroosmosis, yet in some cases this separation technique does not have a replacement alternative. Some attempts to overcome this issue have been performed introducing internal standards into the sample and compensating peak shifting in time. However, existing vector calculation-based methods are computationally intensive for portable instrumentation and usually limited to post-processing applications with 1 or 2 markers. In this work, an original approach of compensating peak migration time shift via signal discretization period correction is proposed. Using the proposed method, the number of reference points or markers that are used for compensation is extended. This method is effective in compensating migration time of peaks in real samples, where high sample injection volumes are used. Using 4 reference peaks in compensation, the method was capable of reducing the relative standard deviation of migration time of the peaks in the electropherograms more than 15 times. Corrected signal discretization periods indicated very high correlations with recorded separation currents, what can be perspective developing an adaptive peak migration time compensation method in capillary electrophoresis.

Keywords: capillary electrophoresis, contactless conductivity detection, peak migration time compensation, real samples, signal discretization period correction

* Corresponding author. Email: audrius.maruska@vdu.lt

INTRODUCTION

Being indispensable in some applications, as a separation technique, capillary electrophoresis suffers from peak shifting and this is a major drawback of separation methods [1, 2]. Even though computational capabilities are increasing and higher performance data acquisition systems are being developed, portable and autonomous analytical instrumentation cannot be utilized with existing peak migration time compensation methods, that are mainly limited to post-processing. Methods for compensation of peak shifting in the electropherograms exist, yet most of them are based on iterative calculation of vectors together with (a) help of an operator, or (b) computationally intensive peak detection algorithms that usually do not provide required performance [3–5]. Even though adaptive methods for sensitivity enhancement or baseline fluctuation compensation have been developed, no adaptive method for peak migration time compensation in the electropherograms has been designed yet [6, 7]. Adaptivity, that is a benefit for a portable instrument, in the method can only be achieved if an independent variable is observed, in parallel predicting a dependent variable of importance. There is no existing peak migration time correction method that shows a significant correlation between independent variable (separation current, temperature, etc.) and peak migration times. Portable instrumentation has been one of the main instrumental analysis development trends during past decades [8, 9]. Now it has been developed into the field of autonomization aimed at investigating remote places on the Earth, Mars and even the Moons of Jupiter and Saturn [10–14]. The requirements for developing such instrumentation are very strict, but the most difficult requirement is that a human operator should not interfere in the analytical process [15]. It can only be achieved if novel mathematical and computerized methods were used. Additionally, separation science is diving into machine learning and artificial intelligence fields which greatly improve the performance and usability of separation methods [16–18]. One of successful applications was the improved quantification of unresolved peaks using artificial neural networks [19]. On the contrary, identifying shifted peaks in a time series, such as electropherogram, is still an issue for any method. Since automatic peak recognition has no good solution and proposed existing algorithms

work well with large datasets of similar chromatograms and electropherograms, its use is limited only to a known beforehand electropherogram pattern, reducing the level of peak misalignment.

The aim of this work was to develop a multi-reference-based signal discretization period correction method for electrophoretic peak migration time compensation and clarify the relationship between a separation current and a corrected signal discretization period.

EXPERIMENTAL

Chemicals

Acetic acid (99.8%) and sodium hydroxide (NaOH) (>99%) were purchased from Rechem (Slovakia). Methanol (MeOH) (99.9%, HPLC grade) was purchased from Sigma-Aldrich (Germany). Bidistilled water was produced in the laboratory using a Fis-treem Cyclon bidistillator (UK).

Sample preparation

The black currant C_2 generation seedlings of a different ploidy level from the seeds variety 'Titania' were grown in the breeding plot of the Institute of Horticulture of the Lithuanian Research Centre for Agriculture and Forestry. Polyploidy of the blackcurrants was induced using the previously published methodology [20]. Only the leaves and buds were sampled in 2017 on the following dates: 22 April, 9 May, 27 May, 5 July and 12 July. The sampled raw material was dried and ground before extraction. For the extraction 0.5 g of ground material was added to an extraction bottle, 10 ml of 75% MeOH was added and the bottles with the extraction content were shaken for 24 h at 200 rpm at ambient temperature $\sim 22^\circ\text{C}$. After the extraction the supernatant was filtered using a paper filter, then 50 μl of supernatant and 50 μl bidistilled water were poured into an analytical vial and the mixture was used for direct injection in a CE instrument (HP3DCE, Hewlett Packard, Germany).

Capillary electrophoresis procedures

Electrophoretic separations were performed according to the previously published procedures [16, 17, 21]. The separation capillary was of 60 cm total length (L_{tot}) and 48 cm effective length (L_{eff}), the inner diameter (I.D.) was 50 μm , the outer diameter (O.D.) was 365 μm . The separation potential applied was +13 kV, temperature 30°C , background electrolyte

(BGE) was 0.5 M acetic acid solution, pH 2.53, hydrodynamic injection was performed at 50 mbar \times 40 s. Before each analysis, the capillary was washed with 0.1 M NaOH solution for 3 minutes, 2 consecutive vials of bidistilled water for 1 min and 3 consecutive vials of BGE for 1 min. After the washing procedure, the electro-conditioning was performed at +13 kV for 10 min.

Detection was performed using a previously developed contactless conductivity detector [22, 23]. Detection conditions were the following: excitation voltage 3.3 V, 32 kHz square wave, open tubular stainless steel electrodes, length 20 mm, I.D. 0.4 mm, O.D. 0.6 mm, detection gap 0.2 mm. The sensitivity of the detector was enhanced using the previously developed migration time-adaptive moving average method [6].

Data acquisition and data analysis was performed using the previously developed software Viewer [22, 23], which was programmed using the open-source software Processing (www.processing.org).

Peak compensation

The data analysis software was upgraded for processing multiple electropherograms, avoiding the manual selection of peaks using Rstudio [24]. Rstudio was also used for the calculation of correlation coefficients between the time series of recorded current and the time series of corrected discretization periods. Peaks were detected using the software developed in the Python programming language with the Scipy package that is intended for scientific analysis of data series [25]. It allowed all analysed electropherograms in a single sequence to be compensated. The method is based on compressing or stretching the electropherogram that is described by the following mathematical statements. The correction coefficient for the 1st segment signal discretization period C_1 (from the start of analysis to the first reference peak) was calculated (Eq. 1):

$$C_1 = \frac{\mu_{t1}}{t1_{e1}}. \quad (1)$$

Here μ_{t1} is the statistical mean of the migration times of the first reference peak for all electropherograms, and $t1_{e1}$ is the migration time of the first reference peak in the corresponding electropherogram. The correction coefficient for the 2nd segment signal discretization period C_2 (from the first reference

peak to the second reference peak) was calculated (Eq. 2):

$$C_2 = \frac{\mu_{t2} - \mu_{t1}}{t2_{e1} - t1_{e1}}. \quad (2)$$

Here μ_{t1} is the statistical mean of the migration times of the first reference peak for all electropherograms, μ_{t2} is the statistical mean of the migration times of the second reference peak for all electropherograms, $t1_{e1}$ is the migration time of the first reference peak in the corresponding electropherogram, and $t2_{e1}$ is the migration time of the second reference peak in the corresponding electropherogram.

The procedure of time scale compensation of the electropherograms is based on that at the given moment t ($t = \Delta \times j$, Δ is signal discretization period (0.216 s), j is data point number), the signal discretization period Δ is adjusted and corrected signal discretization periods Δ_{cn} are obtained (Eq. 3):

$$\Delta_{Cn} = C_n \Delta, n = 1, 2 \dots N. \quad (3)$$

Here C_n is the correction coefficient of the n^{th} segment (from the $n^{\text{th}-1}$ reference peak to the n^{th} peak).

Corrected time values are calculated following Eq. 4):

$$t_{C1}(j) = \sum_{j=1}^{m_1} \Delta_{C1}, j = 1, 2 \dots m_1. \quad (4)$$

Here $t_{C1}(j)$ is the corrected time value in the time series (electropherogram) at data point j ($j = 1, 2 \dots J$) to the 1st reference peak, Δ_{C1} is the corrected signal discretization period of the 1st segment, and m_1 is the data point number at the 1st reference peak migration time. According to the procedure, the corrected time value at data point j is equal to the sum of corrected signal discretization periods from the data point 1 to the data point m_1 . Corrected time values between 1st and 2nd reference peaks are found (Eq. 5):

$$t_{C2}(j) = t_{C1\max} + \sum_{j=m_1+1}^{m_2} \Delta_{C2}, j = m_1 + 1, m_1 + 2 \dots m_2. \quad (5)$$

Here $t_{C2}(j)$ is the corrected time values in the time series (electropherogram) at data point j between 1st and 2nd reference peaks, $t_{C1\max}$ is the maximum corrected time value in the time series (electropherogram)

from the beginning to the 1st reference peak, Δ_{C_2} is the corrected signal discretization period of the 2nd segment, and m_2 is the data point number at 2nd reference peak migration time. According to the procedure, the corrected time values at data point j are equal to the sum of corrected signal discretization periods between 1st and 2nd reference peaks (between m_1 and m_2 data points) and t_{C1max} .

RESULTS AND DISCUSSION

Peak migration time compensation method development

The method development has been performed using real samples instead of modelled analytical standards. The real samples are more complex and cover more investigated conditions. This was done in order to avoid conditions where the optimized method with modelled analytes is not suitable for applications with the real samples. In this case, samples were the extracts (75% MeOH/water) of blackcurrant leaves and buds.

The conditions for electrophoretic separation were selected so that as many organic and inorganic cations as possible would be visible in the electropherogram. The analyses in acidic BGE were lengthy, therefore the capillary flushing procedure with 0.1 M NaOH (pH ~13) was used for the induction of electroosmotic flow (EOF) [26]. Obviously, the washing procedure with an alkaline solution and performing separation in acidic BGE of 0.5 M acetic acid (pH 2.53) together with high injection volumes of real samples decrease the repeatability of migration time of the peaks significantly.

In the real samples the potassium peak is easily identified due to the fact that it usually shows up first [27, 28]. Another peak that can be used as a reference is the peak of the EOF. Using a contactless conductivity detector, the peak of EOF is the highest negative peak: the conductivity of the sample zone with no dissociated analytes is the lowest. Later, it was decided to use a potassium ion and EOF peak as two reference peaks (Fig. 1). Peak migration times in a numeric format can hardly be used for qualitative investigations due to possibility to confuse the peaks (and substances). On the other hand, the peak profile for an analytical chemist is understandable, and it is possible to follow which peak and how much changed the position if the samples are of a similar origin (Fig. 1).

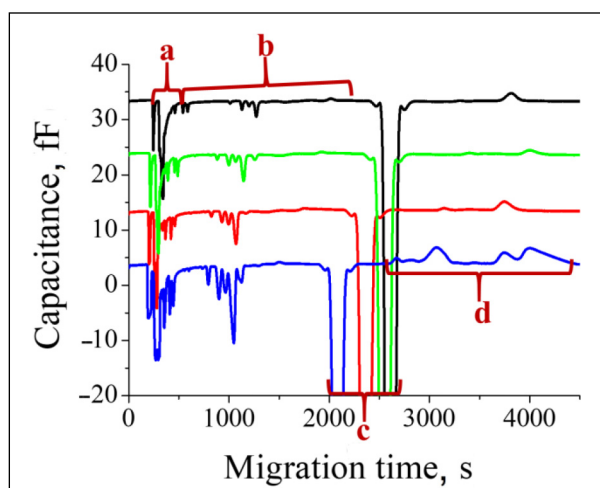


Fig. 1. Representation of peak shifting when real samples are electrophoretically separated. Peaks: (a) inorganic cations, (b) organic cations, (c) EOF, (d) anions

For each electropherogram the K^+ peaks were integrated, peak migration times were found and the mean of migration times was calculated. The mean of the migration time of the reference peak was the statistical position, with the highest possibility of finding it. The procedure was also repeated for the EOF peak. Using the dataset of 32 analyses of different extracts of blackcurrant, the means of K^+ and EOF peak migration times were the following: (i) μ_{tmK^+} was $208.0 \text{ s} \pm 7.2\%$ and (ii) μ_{tmEOF} was $2172.0 \text{ s} \pm 10.7\%$. The reference peak distribution over 32 analyses is represented in Fig. 2. In case K^+ and EOF peaks are not present in the electropherogram, it is possible to use other peaks that exist in all electropherograms or added internal standards.

It was observed that distributions do not fit the Gaussian curve, as this can be explained by the fact that at such separation conditions it is hard to achieve the situation where K^+ ions migrate the effective capillary length in less than 180 s and EOF migrate in less than 1600 s, whereas their velocity can be slowed by multiple means which can be controlled or unpredictable. In some cases, a high migration time shift can occur. For example, the EOF migration time maximum value and the minimum value can nearly differ twice, while with the K^+ peak this has not been observed. Additionally, the distributions of EOF and K^+ peaks are different giving a clue of the dynamic separation process.

If the reference peak is known, it is possible to recalculate the signal discretization period at each

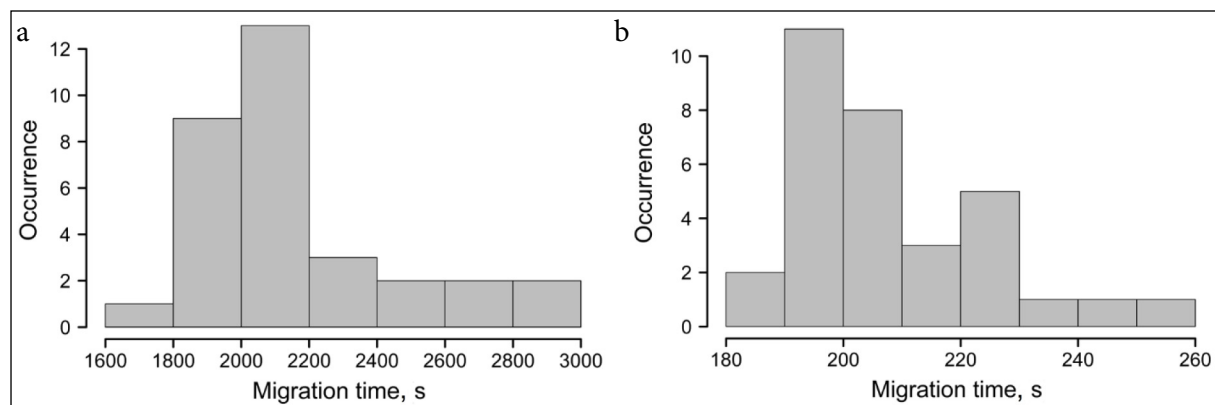


Fig. 2. Reference peak distribution over 32 analyses. (a) Occurrence of EOF peaks in different ranges of migration times. (b) Occurrence of K⁺ peaks in different ranges of migration times

point in the electropherogram and this allows time scale extension or collapsing in the electropherogram. Collapsing and extension depends whether signal discretization correction coefficient values, which are calculated following Eqs. 1 and 2 (in peak compensation section), are higher or lower than 1. Corrected signal discretization periods are calculated following Eq. 3 (in peak compensation section).

The corrected time value in the electropherogram from the start to the 1st reference peak is found following Eq. 4 (in peak compensation section). Corrected time values between 1st and 2nd reference peaks are found following Eq. 5 (in peak compensation section).

The mathematical equations are usually confusing for a chemist when they have to be implemented in practice, therefore the algorithm that is represented in Fig. 3 helps understanding what sequence of calculations has to be performed. The sequence of functions is the following: (i) the number of data points from the start of the analysis to the first reference peak is calculated; (ii) the signal discretization period correction coefficient is calculated for the first compensation segment; (iii) the signal discretization period correction coefficient is calculated for the n^{th} compensation segment ($n = 1, 2 \dots N$, data point number $j = 1, 2 \dots J$, $N = J-1$); (iv) it is checked if all segments have been processed: (a) if no, the steps (iii)–(iv) are repeated, (b) if yes, it is proceeded to the next step; (v) corrected time values are calculated for each data point; (vi) it is checked if all segments have been processed: (a) if no, the step (v) is repeated, (b) if yes, the procedure is terminated.

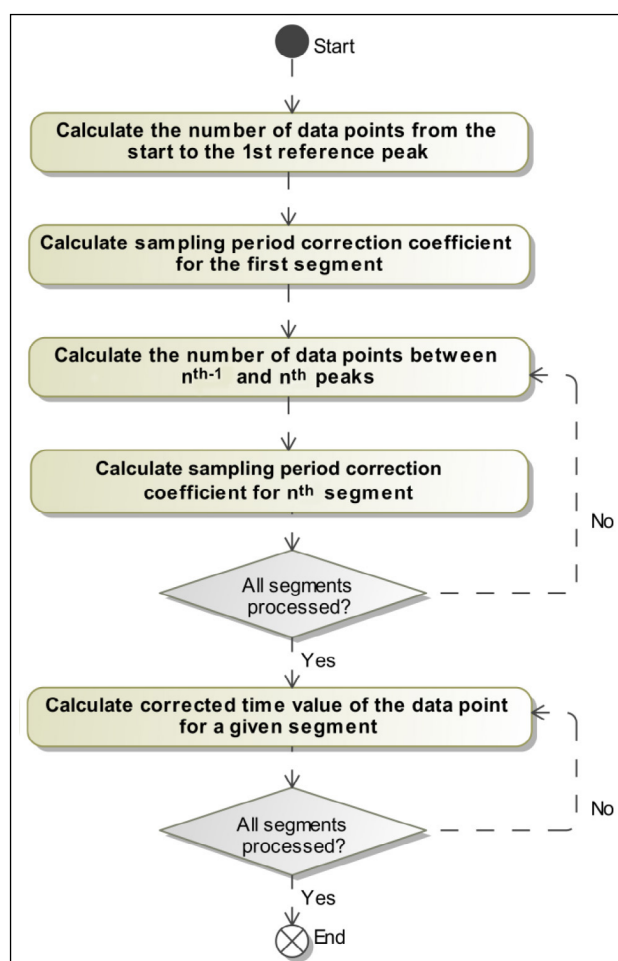


Fig. 3. The algorithm of peak migration time compensation method

The result of this algorithm is the peak migration time compensated electropherogram (Fig. 4e, f). Virtually, any number not bigger than time-series itself of reference points can be added for the segmented time-scale compensation of peaks. The segmentation approach clarifying antiviral properties of medicinal plant extracts has already been reported

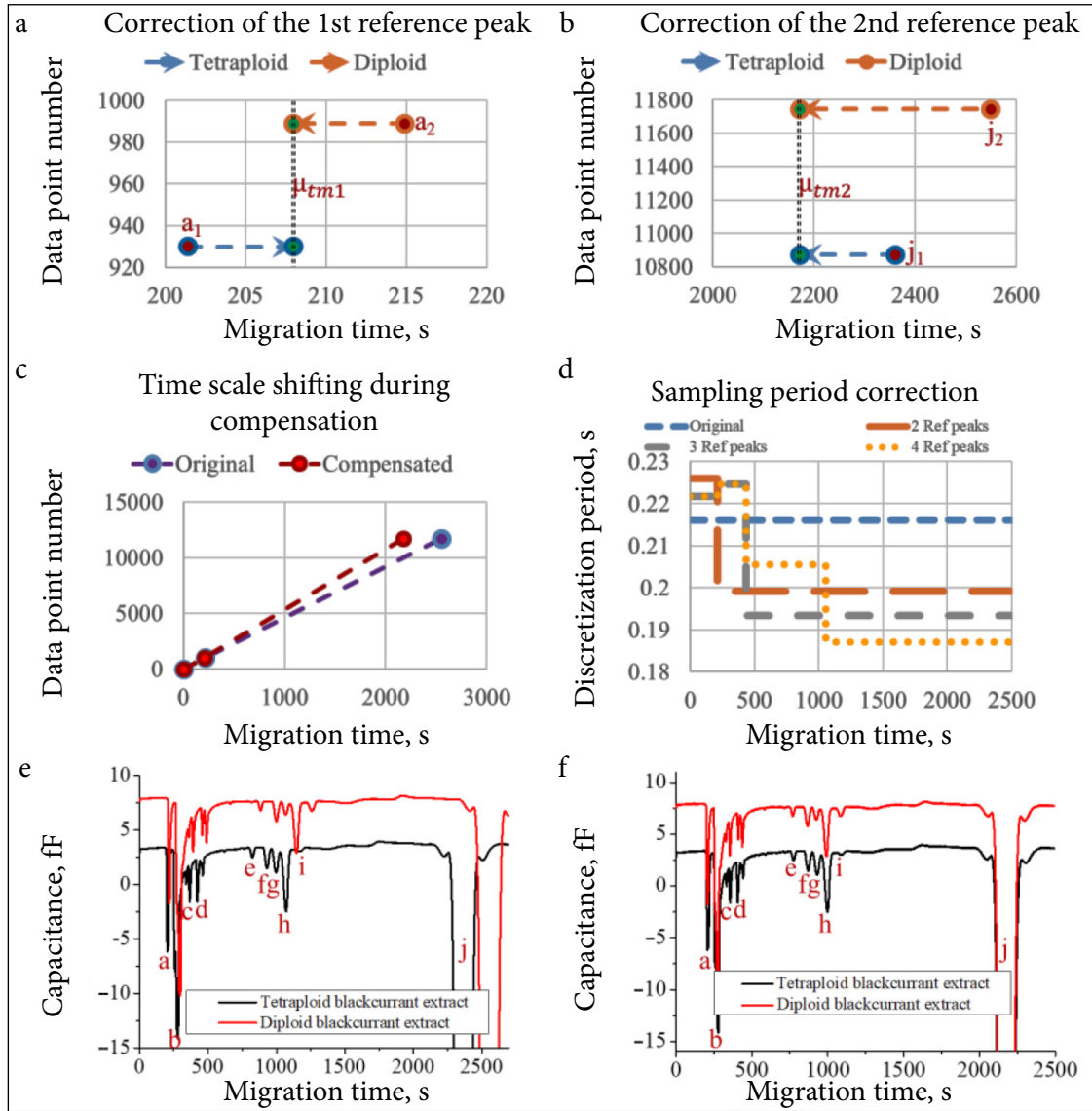


Fig. 4. Representation of the compensation mechanism and the result. (a) Correction of the 1st reference peak (K^+ ions) for the electropherograms of tetraploid and diploid blackcurrant leaves extract, (b) correction of the 2nd reference peak (EOF) for the electropherograms of tetraploid and diploid blackcurrant leaves extract, (c) time scale shifting during the compensation of migration times for tetraploid blackcurrant leaves extract, (d) comparison of the signal discretization period change between the original and differently compensated electropherogram, (E) not compensated electropherogram (tetraploid and diploid blackcurrant leaves extract), (F) compensated electropherogram of tetraploid and diploid blackcurrant leaves extract using 2 reference peaks. Indications: (μ_{tm1}) statistical mean of migration time of the first reference peak, (μ_{tm2}) statistical mean of migration time of the second reference peak (a) K^+ , (a_1) K^+ in the tetraploid blackcurrant extract, (a_2) K^+ in the diploid blackcurrant extract (b) Na^+ , (c), (d), (e), (f), (g), (h), (i) unidentified peaks, used for determination of peak migration shifting, (j) EOF peak, (j_1) EOF in the tetraploid blackcurrant extract, (j_2) EOF in the diploid blackcurrant extract.

[16–18]. The corrected signal discretization period between the n^{th} and the $n^{\text{th}-1}$ reference peaks can be calculated (Eq. 6):

$$C_{n^{\text{th}}} = \frac{\mu_{t_{n^{\text{th}}}} - \mu_{t_{n^{\text{th}-1}}}}{t_{n^{\text{th}}} - t_{n^{\text{th}-1}}}. \quad (6)$$

Here $C_{n^{\text{th}}}$ is the correction coefficient for the signal discretization period between the $n^{\text{th}-1}$ and the n^{th}

reference peaks, $C_{t_{m_{n^{\text{th}-1}}}}$ is the statistical mean of migration times of the $n^{\text{th}-1}$ reference peak for all electropherograms, $\mu_{t_{m_{n^{\text{th}}}}}$ is the statistical mean of migration times of the n^{th} reference peak for all electropherograms, $C_{n^{\text{th}}}$ is the migration time for a given time-series (electropherogram) of the n^{th} reference peak, and $t_{n^{\text{th}-1}}$ is the migration time for a given time-series (electropherogram) of the $n^{\text{th}-1}$ reference peak.

Corrected time values between $n^{\text{th}-1}$ and n^{th} reference peaks are found (Eq. 7):

$$t_{Cn^{\text{th}}}(j) = t_{Cn^{\text{th}}-\text{max}} + \sum_{i=m_{n^{\text{th}-1}}+1}^{m_{n^{\text{th}}}} \Delta_{Cn},$$

$$j = m_{n^{\text{th}-1}} + 1, m_{n^{\text{th}-1}} + 2 \dots m_{n^{\text{th}}}. \quad (7)$$

Here $t_{Cn^{\text{th}}}(j)$ is the corrected time values in the time series (electropherogram) at data point j between $n^{\text{th}-1}$ and n^{th} reference peaks, $t_{Cn^{\text{th}}-\text{max}}$ is the maximum corrected time values in the time series (electropherogram) between $n^{\text{th}-2}$ and $n^{\text{th}-1}$ reference peaks, and $t_{Cn^{\text{th}}}(j)$ is the corrected signal discretization period at data point j . According to the procedure, the corrected time values at data point j are equal to the sum of corrected signal discretization periods between $n^{\text{th}-1}$ and n^{th} reference peaks (between $m_{n^{\text{th}-1}}$ and $m_{n^{\text{th}}}$ data points) and $t_{Cn^{\text{th}}-\text{max}}$.

In Fig. 4 it is visible that the peaks, which showed different migration times, now show similar migration times. Not only reference peaks (K^+ and EOF) but also unidentified peaks show similar positions. Such technique has the element of dynamic programming – the markers that shifted position have been compensated. Dynamic programming is one of the key techniques in artificial intelligence methods used for data processing, allowing to shift the position of markers of importance in voice recognition systems [29, 30].

In Fig. 4a, b it is visible that the first reference peak (K^+) in different electropherograms (tetraploid and diploid blackcurrant extracts) has to be shifted to different directions: (a_1) migrated earlier than the statistical mean and (a_2) migrated later than the statistical mean. The second reference peak (EOF) in both samples migrated later than the statistical mean. During compensation, the peak position corresponding to certain data points must not change, only the time scale. Such observation is an indication that the electrophoretic migration process is dynamic and is different for different samples. The protocols of capillary zone electrophoresis suggest that the samples should be composed of the same solvent containing similar amounts of salts having similar ionic strengths. For real samples this is often not possible and diluting the samples with buffers, or BGE prevents it gaining high pre-concentrations due to the fact that electro-stacking effects can hardly be utilized in high salinity (conductivity) samples [21].

In Fig. 4c, the time change of different data points for original and compensated electropherograms is represented. In Fig. 4d, the signal discretization period change over time is represented for the same electropherogram and using compensation with a different number of reference peaks. In case 2 reference peaks are used for compensation, a single break point at K^+ migration time is obtained and for a given analysis, the signal discretization period has to be reduced, meaning that the total analysis time has to be made shorter. Interestingly, the method of 3 and 4 reference peaks proposes that the signal discretization period from the beginning to the migration time of the K^+ ion has to be increased from 0.222 up to 0.224 s. This means that K^+ ions migrated at a higher velocity for a given electropherogram comparing it to the statistical mean. After the migration time of K^+ ions, the signal discretization period has to be reduced in order to make analysis time shorter. The breakpoints, where the signal discretization period has to be increased and after some time it has to be reduced, are an indication of a dynamic electrophoretic process suggesting that in the beginning the analytes for a given electropherogram migrate at a higher velocity and later the velocity is reduced.

CHARACTERIZATION OF THE PROPOSED METHOD

Thirty-two electropherograms were integrated and the position of 10 peaks (a, b, c, d, e, f, g, h, i and j (Fig. 4d, e)) was indicated for compensated and non-compensated electropherograms. It is visible that the compensation significantly increases the peak position repeatability (Table).

Some peaks show a higher repeatability than others. This is due to the fact that electrophoresis is a dynamic process, where the separation current, especially in the beginning of the analysis, is not stable, providing fluctuations to the whole separation process. Additionally, due to differences of the sample composition (organic and inorganic salt content, dilution) not only separation current is fluctuating but also the EOF. Currently existing methods of compensation are mainly focused on 1 or 2 reference peaks, assuming that the electrophoretic process is stable and static and, most importantly, based on computationally intensive vector calculations for portable instrumentation. For affinity capillary electrophoresis, the shifts of mobility were compensated

Table. Peak position repeatability for not compensated and differently compensated real samples

Peak	Peak migration time shift RSD, %			
	Not compensated	Compensated		
		2 ref. peaks	3 ref. peaks	4 ref. peaks
a	7.2	<0.1*	<0.1*	<0.1*
b	6.0	1.5	1.1	0.9
c	9.0	3.0	1.0	1.0
d	9.1	3.0	<0.1*	<0.1*
e	8.1	3.1	2.4	0.9
f	8.0	3.2	2.4	0.7
g	7.9	3.3	2.4	0.5
h	7.8	3.3	2.5	<0.1*
i	6.3	3.4	2.7	1.1
j	10.7	<0.1*	<0.1*	<0.1*

* Indication of a reference peak.

using correction factors [31]. In a classical compensation approach, for calculating a vector (migration velocity or mobility) the effective length of capillary should be divided by the migration time.

This has to be done iteratively for all data points in the electropherogram. Division is the most computation-intensive operation of the basic mathematical operations (addition, subtraction, multiplication and division). Division of each data point in high-speed (discretization period <0.001 s) analytical systems can be a challenge for low performance microcontrollers such as Arduino. The proposed method is simpler: division is only used calculating the correction coefficient and all the rest operations are the multiplications, or summations. The developed method proposes a new approach of using correction of signal discretization periods between each data point, meaning that the signal discretization period has to be calculated at the reference peaks, used in the whole electropherogram, and the number of such calculations is only limited to the selected number of the reference peaks. The obtained time-series of discretization periods (s) for the 4 reference peaks-compensated electropherogram were correlated with the time series of the recorded separation current (μA) (Fig. 5).

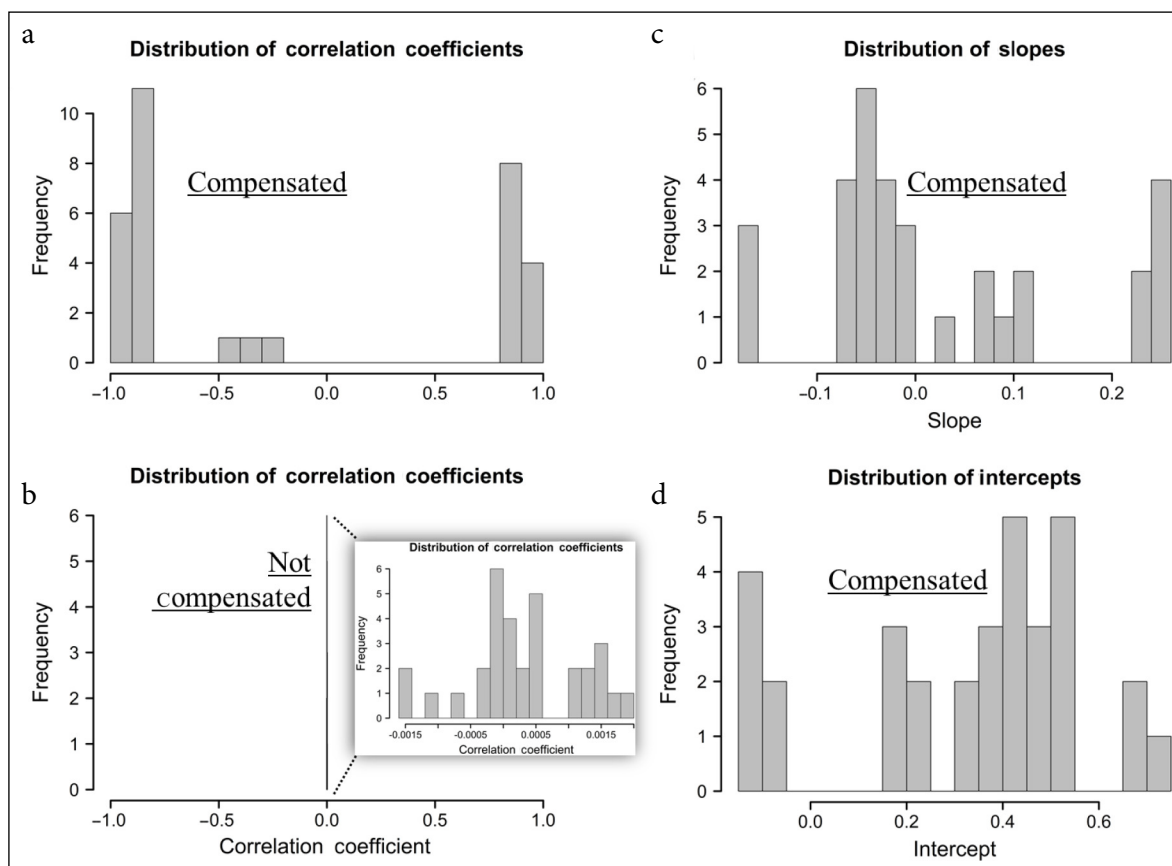


Fig. 5. Representation distributions of the correlation coefficients between the time-series of discretization periods and the recorded separation current for (a) compensated electropherograms, (b) not compensated electropherograms and distribution of (c) slopes and (d) intercepts of compensated electropherograms

Interestingly, very high correlations (absolute values >0.85 , except 1 outlier indicating a low correlation and 2 indicating a moderate correlation) were observed between the compensated time-series of discretization periods and the recorded separation current (Fig. 5a), whereas not compensated electropherograms indicated the correlations between -0.002 and 0.002 (Fig. 5b).

Clearly, the dependency between the separation current and the discretization period of 4 reference peaks compensated electropherograms is high. These findings suggest that an adaptive model predicting the correction of a discretization period by observing the change of current in real-time applications can be developed. Additionally, the slopes and intercepts of linear regression models represented in Fig. 5c, d indicate that the model is complex, both (a) positive and (b) negative slopes and intercepts were observed.

DISCUSSION

In the future, based on findings in this research, a real-time peak migration time compensation method will be developed. This can be achieved via (a) programmed algorithm, that needs deeper understanding between a separation current and a discretization period model, or (b) machine learning/artificial intelligence methods capable of handling time-series. In the main artificial intelligence method group – neural networks, the developed model is based on linear sections that form a separating surface in a multidimensional space [32]. In this work, the compensation method is also based on linear segments that warp the time scale according to the references. In theory, simple linear models rarely explain real samples and observations well due to the fact that in real samples a more sophisticated dependency is present. On the other hand, segmented models of neural networks and deep neural networks are used for very complex modeling not only in a 2-dimensional space, but 3, 4 and higher dimensional spaces. As observed by the results (Table), the compensation using 4 reference peaks outperforms the compensation of 2 reference peaks. These conditions form 4 segments between the start of analysis and the last reference segment. As observed in our work, at the beginning the separation process is slower and later it becomes faster, and this can be explained by the fact that during

different moments of time apparent electrophoretic mobility of the analytes and EOF slightly differs.

This work is an intermediate part that cannot be skipped developing the adaptive real-time peak migration time compensation method.

CONCLUSIONS

In this study, multi-reference peaks-based signal discretization period correction method was developed. The method compensated peak migration time shifts. It reduced the relative standard deviation of peak migration time shift in real samples up to 15.8, 5.4 and 4 times if 4 reference, 3 reference and 2 reference compensation are used. The method is based on an original approach of the signal discretization period correction between different data points, therefore less computational intensity is required for utilizing the method. Very high correlations between time series of the corrected discretization periods and recorded separation current were observed. The method is aimed at use in portable and autonomous instrumentation, where computational intensity is limited, instrumentation requires adaptive calculations and is operated without human interference.

ACKNOWLEDGEMENTS

This research was funded by a Grant (No. 09.3.3-LMT-K-712-02-0202) from the Research Council of Lithuania.

Received 19 May 2020

Accepted 25 May 2020

References

1. E. Farçaş, L. Pochet, J. Crommen, A.-C. Servais, M. Fillet, *J. Pharm. Biomed. Anal.*, **144**, 195 (2017).
2. T. Hirokawa, N. Ikuta, T. Yoshiyama, H. Okamoto, *Electrophoresis*, **22**, 3444 (2001).
3. H. Zhang, H. Chen, *Anal. Methods*, **3**, 745 (2011).
4. R. Stewart, A. Wee, D. B. Grayden, Y. Zhu, in: *Biomedical Applications of Micro- and Nanoengineering IV and Complex Systems* (2008).
5. R. Stewart, I. Gideoni, Y. Zhu, in: *Systems and Computational Biology – Bioinformatics and Computational Modeling* (2011).
6. T. Drevinskas, L. Telksnys, A. Maruška, J. Gorbatsova, M. Kaljurand, *Anal. Chem.*, **90**, 6773 (2018).
7. T. Drevinskas, L. Telksnys, A. Maruška, J. Gorbatsova, M. Kaljurand, *Electrophoresis*, **39**, 2877 (2018).

8. W. R. de Araujo, T. M. G. Cardoso, R. G. da Rocha, et al., *Anal. Chim. Acta*, **1034**, 1 (2018).
9. S. Sharma, L. T. Tolley, H. D. Tolley, A. Plistil, S. D. Stearns, M. L. Lee, *J. Chromatogr. A*, **1421**, 38 (2015).
10. K. R. Elkin, *J. Chromatogr. A*, **1352**, 38 (2014).
11. M. F. Mora, A. M. Stockton, P. A. Willis, *Electrophoresis*, **33**, 2624 (2012).
12. F. Postberg, N. Khawaja, B. Abel, et al., *Nature*, **558**, 564 (2018).
13. M. S. Ferreira Santos, T. G. Cordeiro, A. C. Noell, C. D. Garcia, M. F. Mora, *Electrophoresis*, **39**, 2890 (2018).
14. B. Yang, M. Zhang, T. Kanyanee, B. N. Stamos, P. K. Dasgupta, *Anal. Chem.*, **86**, 11554 (2014).
15. M. Kaljurand, *Trends Environ. Anal. Chem.*, **1**, e2 (2014).
16. T. Drevinskas, R. Mickiene, A. Maruška, et al., *Anal. Methods*, **10**, 1875 (2018).
17. T. Drevinskas, R. Mickienė, A. Maruška, et al., *Chemija*, **29**, 124 (2018).
18. T. Drevinskas, A. Maruška, L. Telksnys, et al., *Anal. Chem.*, **91**, 1080 (2019).
19. R. M. Latorre, S. Hernández-Cassou, J. Saurina, *J. Sep. Sci.*, **24**, 427 (2001).
20. V. Stanys, G. Staniene, T. Šikšnianas, *Acta Univ. Latv. Biol.*, **676**, 235 (2004).
21. T. Drevinskas, M. Stankevičius, K. Bimbiraitė-Survilienė, et al., *Electrophoresis*, **39**, 2425 (2018).
22. T. Drevinskas, M. Kaljurand, A. Maruška, *Electrophoresis*, **35**, 2401 (2014).
23. T. Drevinskas, A. Maruška, V. Briedis, *Electrophoresis*, **36**, 292 (2015).
24. *RStudio Team, RStudio: Integrated Development for R*. (2015).
25. E. Jones, T. Oliphant, P. Peterson, *SciPy: Open Source Scientific Tools for Python* (2001).
26. T. Drevinskas, V. Bartkuvienė, A. Maruška, *Chemija*, **25**, 206 (2014).
27. P. Kubáň, P. Kubáň, V. Kubáň, *J. Chromatogr. A*, **836**, 75 (1999).
28. P. Kubáň, B. Karlberg, P. Kubáň, V. Kubáň, *J. Chromatogr. A*, **964**, 227 (2002).
29. J. Kapočiūtė-Dzikiene, G. Raškinis, *Inf. Technol. Control*, **41**, 318 (2012).
30. R. Lileikytė, L. Telksnys, *Inf. Technol. Control*, **41**, 268 (2012).
31. J. Østergaard, H. Jensen, R. Holm, *J. Sep. Sci.*, **32**, 1712 (2009).
32. D. Taylor, A. Harrison, D. Powers, *Forensic Sci. Int. Genet.*, **30**, 114 (2017).

Tomas Drevinskas, Audrius Maruška, Gintarė Naujokaitytė, Laimutis Telksnys, Mihkel Kaljurand, Vidmantas Stanys, John Cowles, Jelena Gorbatsova

ADAPTYVAUS SMAILIŲ MIGRACIJOS LAIKO KOMPENSAVIMO METODO LINK: DISKRETIZACIJOS PERIODAS ELEKTROFEROGRAMOSE

Santrauka

Kapiliarinėje elektroforezėje dažnai yra stebimas žemas smailių migracijos laiko pakartojamumas dėl elektroosmozės pokyčių, tačiau kai kuriais atvejais nėra alternatyvaus kapiliarinei elektroforezei metodo. Literatūroje gausu aprašytų bandymų, skirtų kompensuoti smailių migracijos laiko pokytį į bandinį įdedant vidinį standartą. Vektorių skaičiavimo principu pagrįsti metodai reikalauja skaičiavimo resursų, kurių nešiojamoji įranga neturi, todėl kompensavimas viena arba dviem žymėmis atliekamas atskira įranga po to, kai duomenys jau yra surinkti. Šiame darbe siūlomas originalus smailių migracijos laiko kompensavimo būdas keičiant diskretizacijos periodą. Naudojant šį metodą, galima pritaikyti daugiau atraminių taškų kompensavimui nei dabar įprasta. Šis metodas yra efektyvus kompensuojant elektroferogramas, kurios gautos tiriant realius mėginius, kai bandinio injekcijos tūris yra santykinai didelis. Pritaikius kompensavimą su keturiomis atraminėmis smailėmis, elektroferogramų smailių migracijos laiko standartinis nuokrypis sumažėjo daugiau nei 15 kartų. Pastebėta, kad pakoreguoti diskretizacijos periodai stipriai koreliuoja su kapiliarinės elektroforezės metu užrašoma skirstymo srove. Tai yra perspektyvu plėtojant adaptyvius smailių migracijos kompensavimo metodus kapiliarinėje elektroforezėje.