MCR-ALS of voltammetric data for the study of environmentally relevant substances

Núria Serrano¹,², Clara Pérez-Ràfols¹,³, Cristina Ariño¹,², Miquel Esteban¹,² and José Manuel Díaz-Cruz¹,²*

¹ Department of Chemical Engineering and Analytical Chemistry, University of Barcelona. Martí i Franquès 1-11, E08028 Barcelona, Spain.
² Water Research Institute (IdRA) of the University of Barcelona.
³ Department of Chemistry, School of Engineering Science in Chemistry, Biochemistry and Health, KTH Royal Institute of Technology, Teknikringen 30, SE-10044 Stockholm, Sweden.

* Corresponding author: josemanuel.diaz@ub.edu; Tel.: +34-93-402-1796

Abstract

A critical revision is made on the main approaches and results arising from the combination of multivariate curve resolution by alternating least squares (MCR-ALS) and electroanalytical measurements in the field of environmental analytical chemistry. Although most of the work done has been focused on the study of the metal binding properties of metal bioregulators such as metallothioneins or phytochelatins, new perspectives appear in the evolving world of sensors for environmental monitoring.

Keywords: Multivariate curve resolution by alternating least squares (MCR-ALS), electroanalysis, voltammetry, metal bioregulators, metal speciation, electrochemical sensors.
1. Introduction

Multivariate curve resolution by alternating least squares (MCR-ALS) is a strategy developed by R. Tauler et al. in the beginning of the 1990s [1, 2]. Formally, it can be written in a similar way as principal component analysis (PCA):

\[ R = C S^T + E \] (1)

where the superscript ‘\(T\)’ means transposed matrix and \(R\) is a multivariate data matrix containing instrumental responses, e.g. a series of spectra measured for different samples (in rows) at different wavelengths (in columns) that is decomposed as the product of a matrix \(C\) related to samples/concentrations and a matrix \(S\) related to wavelengths/pure signals plus an error matrix \(E\) using a certain number of components, which are linear combinations of the original variables. In PCA, \(C\) and \(S\) matrices are named ‘scores’ and ‘loadings’ respectively. They are abstract mathematical entities without any physicochemical meaning, which try to reproduce the essential information contained in the data (the ‘structure’) with the minimum possible number of variables. For this purpose, the NIPALS algorithm extracts one by one a set of new variables, named principal components, from the data matrix [3, 4]. In contrast, MCR-ALS uses a different algorithm (ALS) that, starting from an initial estimation of \(S\) or \(C\) and a prefixed number of components, carries out the iterative calculations:

\[ C = R (S^T)^+ \] (2a)
\[ S^T = C^* R \] (2b)

where the superscript ‘\(^+\)’ denotes pseudoinverse matrix. Calculations according to Eqns. 2a and 2b are repeated in an alternating way until the reproduced \(R_{rep}\) matrix computed as the product \(C S^T\) (Eqn. 1) gets as similar as possible to the experimental data matrix \(R\). A similar couple of equations can be used when the iterations start from an estimation of \(C\) matrix.

The main feature of the ALS algorithm as compared to NIPALS is that it allows the application of restrictions to both \(C\) and \(S\) matrices at each new iteration. Such restrictions are named ‘constraints’ and provide physicochemical information (non-negative values, unimodal signals, mass balances, presence or absence of certain species, equilibrium or kinetic equations...) that confer physicochemical meaning to \(C\) and \(S\). In this way, \(C\) matrix contains the ‘concentration profiles’, i.e., the evolution of the concentration of every component along the different measurements and \(S\) matrix contains the ‘pure signals’, i.e., the spectra of every component alone at unity concentration. Indeed, MCR-ALS can also be understood as an iterative version of classical least squares calibration (CLS) [3, 4] where neither \(C\) nor \(S\) nor the number of components are known.
Although it was originally conceived for the analysis of spectroscopic data, quite soon MCR-ALS found interesting applications in other types of measurements, and nowadays it is widely used in the fields of analytical chemistry and environmental sciences [5-8]. Moreover, the application of MCR-ALS has been stimulated by the development of a graphical user interface (GUI) version of the program and a support website [9-11].

The first application of MCR-ALS to electroanalytical data was done quite early on, in 1995 [12], but, despite the appealing features shown [13-15], MCR-ALS is still less popular in electroanalysis than in other areas of instrumental analysis such as UV-vis, fluorescence and NIR spectrophotometries or even chromatographic techniques. The reasons for that seem to be related to the prevalence of electrochemical models based on fundamental equations (what is called *hard modelling* in opposition to the *soft modelling* concept represented by MCR-ALS) and the frequent non-linearity of electrochemical data, which may hinder the application of MCR-ALS approach [16, 17]. Moreover, when using multivariate analysis, many authors prefer more ‘classical’ approaches such as PCA for data exploration and partial least squares (PLS) for calibration [14].

Nevertheless, MCR-ALS presents especial potentialities as compared to PCA or PLS, mainly arising from the above-mentioned use of constraints to get meaningful concentration profiles and pure signals instead of a set of abstract scores and loadings. It must be pointed out that, in order to get an optimal benefit from the concentration profiles and the pure signals of MCR-ALS, the different samples/experiments constituting the rows of the data matrix should not be placed randomly there. On the contrary, they should correspond to consecutive measurements recorded along a progressive variation of the experimental conditions (*e.g.*, at increasing or decreasing pH values or metal-to-ligand ratios). In this way, the concentration profiles will show the *evolution* of the concentration of the species as a function of the changes made along the experiment. Unlike PCA, which tries to explore data and find differences between samples from different origin, or PLS, which tries to predict the concentrations of many unknown samples by comparison with the (randomly ordered) responses of a set of standards, MCR-ALS tries to describe the evolution of the components of a single sample submitted to a progressive change of experimental parameters. This can be especially useful for the calculation of stoichiometries, kinetic parameters and equilibrium constants.

Additionally, in the case of voltammetric (and chronopotentiometric) measurements, the application of MCR-ALS has some peculiarities:

- The components are indeed associated to electrochemical processes (oxidations, reductions, capacitive phenomena...) and not to chemical species. Although in many situations, one
chemical species generates only one electrochemical process and hence just one signal, quite often one species undergoes more than one process and produces several signals.

Most electrochemical processes produce signals with a characteristic shape (sigmoid, peak, ...) that can be reproduced with a parametric function (e.g., a Gaussian peak). This allows one to apply a restriction called ‘shape constraint’ which forces the pure signal of a component/electrochemical process to follow a prefixed shape by substituting the values of the pure signal obtained at every iteration by the values of the parametric function fitted to them.

Taking into account all these considerations, Figure 1 summarizes a typical application of MCR-ALS to electrochemical data resulting from a voltammetric titration of a metal ion with a ligand or vice versa [16]. The ALS iterations usually start with a set of initial estimations of pure signals (S matrix) obtained from measurements of pure solutions, from the ‘visual’ adjustment of peak-shaped functions to the experimental data set [16, 18], or by using the simple-to-use interactive self-modelling mixture analysis (SIMEFLSMA) approach, which selects the purest variables from the data set [19, 20]. Evolving factor analysis (EFA) [21, 22] can also be used to get initial estimations of either C or S matrices. Then, ALS iterations are carried out as described by Eqns. (2a) and (2b) applying a set of constraints. Once the error matrix E has been minimized, the optimized C and S matrices show the evolution of the concentration of every component as a function of the metal-to-ligand or ligand-to-metal ratio and the pure voltammogram of every component, respectively. As discussed later, these results can be very useful in the estimation of stoichiometries and stability constants of the formed complexes (C matrix) and their relative stability and electrochemical reversibility (S matrix).

However, the physical meaning of MCR-ALS results provided by ‘natural’ constraints is an advantage counterbalanced by a certain ambiguity. For instance, in PCA the factor matrices (scores and loadings) are orthogonal and follow the directions of maximum explained variances and, apart from some scaling or normalization options, the solution is unique for a given number of components. In contrast, the iterative ALS strategy shown in Eqns. 2a and 2b and the use of certain constraints hinder the orthogonality of C and S matrices obtained in MCR-ALS. As a consequence, the solutions achieved by using MCR-ALS are not unique and, therefore, have an unknown amount of ambiguity [23, 24]. Two types of ambiguities are distinguished in MCR-ALS: intensity ambiguities and rotational ambiguities. The first one is present in any factor analysis decomposition unless the scale of one of the two factors matrices is fixed. Fortunately, this problem can be easily fixed in many cases by using closure constraints (e.g., applying mass balance equations). More problematic is rotational ambiguity, which happens when a set of different solutions (and linear combinations of them) fit the
experimental data equally well, although they represent really diverse physico-chemical situations [23, 24]. The best way to reduce rotational ambiguity is by applying appropriate constraints, that can be the usual ones (non-negativity, unimodality ...) or some specifically designed to minimize rotational ambiguity, such as the so-called area correlation constraint, to be used in the analysis of second order data [25]. Another convenient strategy is the analysis of augmented data matrices coming from a set of complementary experiments, especially when the known absence of some components in some parts of the experiments can be used to apply selectivity and/or local range constraints. As shown by Olivieri and Tauler [26], the use of multiple data sets having common constituents drastically reduces rotational ambiguities as compared to the individual analysis of each data set separately. In the case of electroanalytical data, the shape constraint is also a key tool to decrease rotational ambiguity [13].

Although most MCR-ALS users are aware of the presence of rotational ambiguities in their results and try to decrease them, practically nobody estimates their real extent. Of course, this is a difficult task from a theoretical point of view and can be made in different ways. Among these, the approach by Jaumot and Tauler, named MCR-BANDS, is especially versatile [24]. It is implemented in a GUI similar to this of the standard MCR-ALS GUI [10] and provides confidence bands around the concentration profiles and pure signals obtained by MCR-ALS delimiting the range of feasible solutions. The increasing awareness about rotational ambiguities will surely increase the use of this kind of programs in the future. In the same way as the credibility of an univariate result depends on the presence (and the value) of its standard deviation, the credibility of MCR-ALS results will demand soon the support of a realistic estimation of rotational ambiguities.

In the following sections, we will review the most relevant applications of MCR-ALS for environmental analytical purposes.

2. Study of biomolecules involved in heavy metal detoxification

As it was mentioned above, MCR-ALS coupled to voltammetric data can be a powerful tool to study the complexation between metal ions and different ligands. These studies can be particularly relevant for environmental purposes since they can help us to understand the metal binding characteristics of biomolecules involved in the detoxification of cells in the presence of metal stress, which can be then used, for example, to improve bio-remediation treatments. This is the case of peptides like metallothioneins in animals and phytochelatins in plants, but also some related peptides like glutathione. Figure 2 summarizes the structure of some of these molecules, which have in common the presence of thiol groups from the aminoacid cysteine with a strong affinity for heavy
metal ions, mainly Cd(II), Zn(II), Pb(II) and Hg(II). The voltammograms registered in these heavy metal-biomolecule systems present plenty of overlapping peaks corresponding to the reduction of free and bound metal ions and also to redox processes of thiol groups. In many cases, the peaks increased and decreased at the same peak potential along the experiments, suggesting that the electrochemical measurement was faster than the association-dissociation kinetics of metal complexes with biomolecules (electrochemically inert complexes). This ‘stability’ of the signals at a fixed potential was essential to ensure the linear character of the voltammetric data and, therefore, a proper application of MCR-ALS. As for techniques and electrodes, most of the studies were carried out by differential pulse polarography (DPP) with static mercury drop electrode (SMDE) because of the extraordinary reproductibility of the data so obtained, but the toxicity of mercury forced the use of environmentally friendly alternatives such as differential pulse voltammetry (DPV) in bismuth films deposited on glassy carbon and screen-printed electrodes as substrates.

The usual experimental design in this kind of studies was the voltammetric titration of the biomolecule with successive additions of one or more metal ions and vice-versa. Some experiments changing pH values were also made. The analysis of the corresponding data matrices with MCR-ALS produced concentration profiles that, considering the metal-to-ligand ratios at which every species/process appears, stabilizes and disappears, allowed the estimation of metal/biomolecule stoichiometries. MCR-ALS also generated pure voltammograms that were very useful to confirm the identity of the chemical species and the electrochemical processes involved, as well as the relative stability of the metal ions bound to the biomolecule in different ways (usually a higher stability of the metal binding implies a shift of the metal reduction signal to more negative potentials, since reduction becomes more difficult).

Chronologically, the first studies were focused on Cd(II)-glutathione complexes [27, 28] and Cd(II) and Zn(II) binding by metallothioneins and some of their fragments [18, 29, 30]. Later, the Cd(II), Zn(II) and Pb(II) binding properties of small peptides [31, 32] and phytochelatins of different lengths (PC$_2$, PC$_3$, PC$_4$ and PC$_5$) were investigated by DPP [32-38]. Moreover, adsorptive accumulation in constant current stripping (AdSCP) was applied as an alternative for the electrochemical study of metal complexation of glutathione or PC$_2$ with zinc ions on hanging mercury drop electrode (HMDE) [39]. Finally, the complexation processes of Cu(II) by the phytohormone and possible antitumoral agent 6-benzylaminopurine (BAP) were also considered through DPP measurements [40].

Here below, we will discuss in some detail the work in ref. [32] for illustrative purposes. Figure 3a shows the set of voltammograms obtained at a mercury electrode during the voltammetric titration of a solution of the peptide Cys-Gly with a solution of Pb(II)-ion. The application of MCR-ALS
generates the concentration profiles and the pure voltammograms shown in Figures 3b and 3c, respectively. The integrated analysis of both sets of results allowed us to propose the formation of the successive ML and ML₂ complexes shown in Figure 4. This explains the evolution of the species as follows. In the beginning of the titration only the free peptide is present (component 3). As the Pb(II) solution is being added, the peptide starts to bind the metal ion and, consequently, its concentration (component 3) decreases, while two new species appear: the ML complex (component 5) and the ML₂ complex (component 4). With a large excess of metal, the free Pb(II)-ion appears (component 2). Figure 3c supports the proposed stoichiometries in different ways: e.g., the maximum concentration of ML₂ is produced at a 0.5 M/L ratio and the extrapolation of the line corresponding to the free metal (component 2) crosses the x-axis at a M/L ratio of 1. Finally, component 1 is attributed to the oxidation of the mercury electrode favored by the peptide complexes.

In normal situations like that, MCR-ALS could be successfully applied to single matrices typically obtained along a DPP titration with SMDE adding metal ion to the biomolecule or biomolecule to the metal ion. The main constraints used were non-negativity for signals and concentrations, signal shape for voltammograms (usually peak shape) and selectivity for signals and concentrations (e.g., forcing zero concentration of bound metal in the absence of biomolecule or forcing zero current at potentials where the species do not contribute to the voltammogram). Then, the concentration profiles and the pure signals obtained were sufficient to estimate the stoichiometry and the relative stability of the metal complexes formed. Quite frequently, the data corresponding to metal-to-ligand and ligand-to-metal titrations had to be treated simultaneously by means of column-wise augmented data matrices. This methodology was also applied to study the competition between different biomolecules to bind a certain metal ion and to investigate the exchange between different metal ions bound to a biomolecule.

In some cases, however, the excessive overlapping of signals, the progressive shift of the peaks along the potential axis, the changes in the peak shape (e.g. peak broadening) or the losses of linearity hindered the application of MCR-ALS to voltammetric data. In these cases, both the experimental set-up and the chemometric models can be adjusted in order to enable the successful application of MCR-ALS. Some examples of these ‘advanced’ MCR-ALS approaches were:

- Combination of signals obtained by voltammetry and circular dichroism into a row-wise augmented data matrix for an integrated MCR-ALS treatment in order to improve the resolution of the system [41].
- Use of bismuth films coated on glassy carbon and screen-printed electrodes to decrease the interference of the signals generated by the mercury of SMDE in the study of Pb(II) binding by glutathione and phytochelatins [42, 43].

- Use of a gold rotating disk electrode (Au-RDE) to study the Hg(II) binding by phytochelatins [44].

- Use of additional techniques like isothermal titration calorimetry (ITC) or mass spectrometry (MS) to complement the information extracted by MCR-ALS from voltammetric data [37, 44].

- Use of sigmoid shape constraints for the concentration profiles in the MCR-ALS treatment of voltammetric data obtained along pH titrations (experiments where the changing variable was pH thanks to the addition of a basic solution to an initially acidic mixture metal/biomolecule) [45].

- Development of several mathematical approaches to correct data for potential shifts and peak broadening, which decrease their linearity [46-50].

At this point, the complexes of Hg(II) with phytochelatins and related peptides deserve an especial mention. Unlike other heavy metal ions, Hg(II) binds so strongly to these biomolecules that the binding remains unchanged while the sample is passing through a chromatographic column. This made possible the study of synthetic solutions first and natural extracts later by means of liquid chromatography with amperometric detection [51, 52]. In these measurements, detectors typically included glassy carbon disks or carbon-based screen-printed units as working electrodes to carry out the oxidation of both free and bound forms of the peptide. In general terms, the good chromatographic separation of the species considered made unnecessary the use of chemometrics. Nevertheless, the full chromatograms or the areas of selected peaks were used to discriminate samples e.g. by means of PCA [53].

Among the results of such investigations, we could summarize the following ideas:

- Thiol-rich peptides acting as metal bioregulators can bind heavy metal ions with a large diversity of stoichiometries, depending on the number of thiol units per molecule and on the molar ratio between metal ion and peptide.

- Metal ions are bound to these molecules with different degrees of strength, depending on the number of thiol and carboxylic groups contributing to the binding (e.g., a Cd(II)-ion bound to four thiol groups is much more stable than another Cd(II)-ion bound to two thiol and two carboxylate groups). The proportion of each type of metal binding dramatically depends on the proportion of metal and biomolecule (large excesses of the biomolecule favor the strongest binding).
- In this kind of systems, the association-dissociation kinetics of metal binding is usually fast enough to reach chemical equilibrium in solution in a few seconds (the time required for a homogeneous mixing during a titration) but still slower than the voltammetric measurement (electrochemically inert complexes), which avoids the shift of the signals along the potential axis. In contrast, metal ions weakly bound to the biomolecule may have a faster kinetics (closer to electrochemically labile complexes) that could cause potential shifts and, hence, decrease the linearity of the data.

- When metal ions compete to be bound by the same thiol rich based biomolecule, the order of affinity is Cu(II)>Cd(II)> Zn(II)>Pb(II). Metal ions can displace from the thiol rich biomolecule other metal ions with less affinity, which were previously bound, and this happens with a fast kinetics. Depending on the stoichiometries, a thiol rich biomolecule can bind simultaneously different types of metal ions.

- When phytochelatins (PC_n) of different chain length (given by n) compete to bind the same metal ion, longer chains win, although at high n values there is not much difference. Thus, for n between 1 and 3, longer peptides easily displace metal ions from their complexes with shorter ones to produce a more stable binding (signals become more negative), whereas the inverse process is not observed. In contrast, for n between 3 and 5, the signals of Cd–PC_n complexes are very similar, thus suggesting that metal binding is more determined by the nature of the functional groups directly involved in the binding process than by the whole molecule. In any case the formation of mixed complexes with phytochelatins of different lengths were observed [38].

3. Determination and speciation of heavy metal ions and other pollutants

Voltammetric techniques, especially anodic stripping voltammetry (ASV) and adsorptive stripping voltammetry (AdSV), are very convenient for the determination of heavy metal ions in environmental samples [54]. This is not only due to their excellent detection limits, but also because they can provide information about metal speciation, i.e., the distribution of every metal ion into its different physicochemical forms (which can have different degrees of toxicity). Unfortunately, this information is not easy to extract from the data.

A valuable tool to evaluate the speciation of certain heavy metals is to study their complexation ability with relevant ligands present in the environment. MCR-ALS can be useful for this purpose, as it happens in the work by Asadpour-Zeinaly et al., who applied MCR-ALS to determine the stability constants of two successive complexes of Ni(II) and Cu(II) with glycine [55] or in the work by Antunes et al., who studied the binding of Cd(II) by fulvic acids [56].
MCR-ALS is also very efficient for the determination of heavy metal ions producing overlapping signals, as shown by Antunes et al. in the analysis of mixtures of Cd(II), In(III), Pb(II) and Tl(I) by means of ASV at a mercury electrode [57]. Good results were also obtained by Alves et al. by applying MCR-ALS to the determination of Cd(II) and Pb(II) in surface river water samples with the vibrating gold microwire electrode, an environmentally friendly device [58]. When unexpected interferences hinder the accurate determination of a metal ion (or any other pollutant) in the analysis of natural water samples, a good tool to increase the performance of MCR-ALS is the use of second order data (what some authors call ‘the second order advantage’) [59] taking a new electrochemical parameter to consider besides potential and current. For instance, Abdollahi et al. measured currents by DPV as a function of both the potential and the pulse duration and this allowed an improved determination of Pb(II) in river samples [60].

Besides heavy metals, voltammetric methods have also been applied to determine organic pollutants like pesticides. The work by Mora Díez et al. [61], for instance, shows that MCR-ALS can be very useful in the voltammetric determination of ethiofencarb with a glassy carbon electrode in the presence of two other pesticides of the same family, fenobucarb and endiocarb, producing signals strongly overlapped to that of ethiofencarb. For this purpose, the authors take advantage of the alkaline hydrolysis of the pesticides (which generate different phenol derivatives) to get second order data. This research will be described in more detail now, as it is a good example of the strategy based on the ‘second order advantage’. Figure 5 shows that after 4 minutes of alkaline hydrolysis of the three pesticides, their voltammograms (initially very similar) can become quite different as the original compounds disappear and the different derivatives arise. Then, experiments are designed with standard solutions of every substance to obtain voltammograms at four hydrolysis times (1, 2, 3 and 4 min). This produces a matrix like that of Figure 6 for every standard solution, i.e., a set of currents as a function of potential (first order) and hydrolysis time (second order). The matrices obtained for all the solutions can be packaged into a tensor in order to carry out a three-way data analysis such as parallel factor analysis (PARAFAC) or unfolded into a column-wise augmented data matrix to apply MCR-ALS. When this last approach is used, results like these of Figure 7 are obtained. Figure 7a shows the pure voltammograms extracted for every pesticide (S matrix) and Figure 7b the evolution of the concentration of every substance (ethiofencarb in black, fenobucarb in purple and bendiocarb in red) as a function of the hydrolysis time and for every standard solution of ethiofencarb. This allowed the authors to build a calibration model that was successfully validated with test solutions containing all three pesticides.

Different types of drugs [62-64] have also been determined by voltammetric means, mostly in biological and pharmaceutical samples. An example of this is the work by Meshki et al. [65], who
used second order voltammetric data and MCR-ALS to determine the sulfa drugs sulfamethoxazole (SMX) and sulfamethizole (SMT) in human urine and serum or the paper by Koobbi et al. [66], who determined betaxolol and andatenolol in human plasma using a similar methodology. Now that many pharmaceutical products are becoming emerging pollutants in the environment, methods like these could be adapted to the determination of these substances in environmental samples, where they exist in much lower concentrations and with quite complex matrices. Organic UV filters and, in general, personal care products are another group of emerging pollutants susceptible to be determined in environmental samples with electrochemical sensors enhanced by chemometric tools like MCR-ALS [67]. Indeed, chemometric methods in general and MCR-ALS in particular have been extensively used to determine organic pollutants in environmental samples, but the proportion of electroanalytical methods in these studies is still rather small [6].

An important challenge in the analysis of environmental samples is the presence of matrix effects. If the signals of the different analytes are not well resolved, calibration can be problematic. In the case of spectrophotometric measurements, second order information can be helpful to carry out calibration by standard addition, even using replicates of the samples to exploit the second-order advantage with first-order data [68]. In a recent work, our group has shown that the shape constraint characteristic of voltammetric measurements is a key factor to apply MCR-ALS to multivariate standard addition with no need of second order data. Moreover, the proposed method allows one to simultaneously add all the analytes from the same standard solution, which substantially reduces the number of required measurements, and consequently the total analysis time [69]. Figure 8 summarizes this methodology. This is a particular case of the scheme shown in Figure 1 where the first row of the data matrix contains the voltammogram of the original sample and the next rows host the voltammograms measured after successive additions of a standard solution containing all the analytes. A crucial point of this method is the use of a signal shape constraint to discriminate the contribution of each analyte present in the sample and the standard solutions to the overall voltammetric signal. Although the use of asymmetric peak functions can produce better reproductions of the experimental matrix, a simple and fully symmetric function like the Gaussian peak is more reliable due to its much lower trend to ‘transfer’ the signals of minor components to the queues of major components’ peaks. The use of this modality of MCR-ALS generates concentration profiles that can be treated in the same way as univariate signals in the classical standard addition method (just extrapolating to zero y-value). Figure 9 shows an application example with mixtures of the isomers hydroquinone and catechol. As discussed earlier, the consolidation of MCR-ALS as a valid tool to carry out multivariate standard addition could strongly contribute to its
popularization for the voltammetric determination of persistent and emerging pollutants in environmental samples.

4. Conclusions

MCR-ALS possesses excellent features for the treatment of electrochemical data obtained in environmental studies. It is a versatile strategy that allows one to combine data from different sources (data fusion) in augmented matrices and has many constraints available that can take advantage of the peculiar characteristics of electrochemical measurements (especially signal shape). In addition, MCR-ALS is particularly appropriate to follow the evolution of concentration profiles along the experiments, which makes it especially useful for the study of stoichiometries and stability constants as well as to carry out calibrations by standard addition in a very simple and convenient way. Moreover, programs are written in Matlab environment [70] and can be freely accessed and modified to include new data pretreatments and constraints. The main drawback is the need for linear data, which can be neglected if the deviations from linearity are not excessive or can be reasonably corrected with some specific pretreatments. Also, rotational ambiguities can be problematic and should be evaluated by using approaches such as MCR-BANDS.

Nevertheless, and despite such good features, the use of MCR-ALS in the combined field of electroanalysis and environment is still scarce as compared to more popular techniques like PLS or even the application of MCR-ALS to other types of measurements. We hope that in the next years, the need for fast, simple, on site and cheap sensing for environmental control will stimulate the development of electrochemical and hybrid sensors and an extensive application of MCR-ALS for the treatment of the resulting data.

Acknowledgments

This work is supported by the Generalitat of Catalonia (Project 2017SGR311), the Faculty of Chemistry of the University of Barcelona and the Water Research Institute (IdRA) of the University of Barcelona.

References

https://doi.org/10.1002/cem.1180090105.


https://doi.org/10.1080/10408340600970005.

https://doi.org/10.1016/j.talanta.2009.09.044.


https://doi.org/10.1016/j.chemolab.2014.10.003.


https://doi.org/10.1016/0022-0728(95)04015-G.


https://doi.org/10.1021/ac00130a035.


https://doi.org/10.3390/s19153286.

https://doi.org/10.3390/s19184039.


https://doi.org/10.1021/acs.analchem.9b05477.
**Figure Captions**

**Figure 1**
General scheme for the application of MCR-ALS to voltammetric data.

**Figure 2**
Structure of typical molecules involved in the bioregulation of heavy metals.

**Figure 3**
Application of MCR-ALS to the data obtained in the differential pulse voltammetric titration of $2 \times 10^{-5}$ mol L$^{-1}$ of the peptide Cys-Gly with Pb$^{2+}$ at pH 6.8 by using a SMDE. The experimental data matrix is shown in a), whereas b) and c) show the concentration profiles and the pure signals, respectively, obtained with MCR-ALS. The constraints applied were non-negativity for both concentrations and signals, selectivity at the beginning of the titration and signal-shape for the voltammograms of all components except component 4. The components are assigned to the following species/processes: 1) oxidation of the Hg of the electrode favoured by the peptide, 2) free Pb$^{2+}$-ion (M), 3) free peptide (L), 4) 1:2 complex (ML$_2$) and 5) 1:1 complex (ML). Reproduced from ref. [28] with permission.

**Figure 4**
Proposed structures of the complexes ML (a) and ML$_2$ (b) from the analysis of the results shown in Figure 3. Adapted from ref. [28] with permission.

**Figure 5**
Differential pulse voltammograms of the pesticides ethiofencarb, bendiocarb and fenobucarb after 4 min of alkaline hydrolysis. Reproduced from ref. [57] with permission.

**Figure 6**
Two way data set (current versus potential and time) obtained for a validation sample containing ethiofencarb, bendiocarb and fenobucarb. Reproduced from ref. [57] with permission.

**Figure 7**
a) Pure signals retrieved by MCR-ALS from the validation sample data in Figure 6; b) Evolution of the concentrations of ethiofencarb (in black), fenobucarb (in purple) and bendiocarb (in red) as a function of the hydrolysis time, where the validation sample, containing all three pesticides, is
compared with three standard solutions containing only ethofencarb (concentrations 3.60, 6.25 and 10.00 μmol L⁻¹). Reproduced from ref. [57] with permission.

Figure 8
Scheme of multivariate standard addition of voltammetric data by means of MCR-ALS and Gaussian signal shape. Adapted from ref. [65] with permission.

Figure 9
Experimental (black) and calculated (blue) differential pulse voltammograms with corresponding calibration curves (inset) for the simultaneous determination of hydroquinone (HQ) and catechol (CC) in 0.1 mol L⁻¹ phosphate buffer at pH 7 using a graphene screen-printed electrode in a solution containing 4.5 μmol L⁻¹ HQ and CC (a), 9.1 μmol L⁻¹ HQ and 4.6 μmol L⁻¹ of CC (b), and 4.6 μmol L⁻¹ of HQ and 9.1 μmol L⁻¹ of HQ (c). Reproduced from ref. [65] with permission.
Figure 1

Experimental data matrix (R)

Estimation of the number of components:
- Singular value decomposition (SVD)
- Evolving Factor Analysis (EFA)
- Visual inspection of the matrix

Constraints:
- Non-negativity
- Signal shape
- Chemical equilibrium
- Selectivity
- Unimodality
- Mass balance

Initial estimation of C or S:
- EFA (C or S)
- SIMPLISMA (S)
- Pure voltammograms (S)
- Gaussian peaks from the visual inspection of R (S)

Error matrix

Iterative ALS method:
$$S^T = C^T R$$
$$C = R (S^T)^+$$

(+ : pseudoinverse)

Information about:
- stoichiometries
- stability constants

Information about:
- relative stability of complexes
- electrochemical reversibility
Cysteine (cys)

Heavy metal ions:
Cd$^{2+}$, Zn$^{2+}$, Hg$^{2+}$, Pb$^{2+}$...

Phytochelatins, PC$_n$
[$\gamma$-glu-cys]$_n$-gly, $n=2...11$

Glutathione, GSH
(\gamma-glucysgly)

Metallothioneins, MT

Domain $\beta$

Domain $\alpha$
Figure 3

(a) 3D graph showing the relationship between -I/nA and E/V vs. [Pb]/[Cys-Gly].

(b) Graph showing normalized pure currents vs. E/V.

(c) Graph showing concentration vs. [Pb]/[Cys-Gly].
experimental data / reproduced data

peakmaker

constrains:
- non-negativity
- signal shape (gaussian)

initial values of $S^T$ from visual fitting of gaussian peaks

$n_C$ components

$n_E$ potentials

$n_V$ voltmograms

$n_Y$ standard additions

$n_C$ components

$n_E$ potentials

concentration profiles

pure signals

extrapolation to zero to obtain the concentrations in the original sample

Figure 8
Figure 9

(a) 1:1 (HQ:CC)

(b) 2:1 (HQ:CC)

(c) 1:2 (HQ:CC)

\[ R^2_{CC} = 0.997 \]
\[ R^2_{HQ} = 0.998 \]

\[ R^2_{CC} = 1.000 \]
\[ R^2_{HQ} = 0.999 \]