An analysis of distinguishing composite dissolved metal–ligand systems measurable by stripping voltammetry

Cédric Garnier, Ivanka Pižeta, Stéphane Mounier, Vlado Cuculić, Jean Yves Benaim

PROTEE Laboratory, University of Sud Toulon Var-BP 20132, 83957 La Garde, France
Center for Marine and Environmental Research, Ruder Bošković Institute, P.O. Box 180, Bijenička c. 54, 10002 Zagreb, Croatia

Received 7 October 2004; received in revised form 17 December 2004; accepted 16 February 2005
Available online 14 March 2005

Abstract

Voltammetry is a method able to distinguish in certain degree the speciation of dissolved metals. An analysis of its ability to discern composite and more complex dissolved metal–ligand systems has been carried out by simulating the experiments for determination of metal–ligand complexing parameters. Logarithmic equidistant addition of metal was presumed, covering 2.5 decades. The data obtained with the preset parameter values were subsequently fitted to the presumed models. Data points under the detection limit DL = 10^{-10} mol L^{-1} were eliminated and random noise following a realistic shape was added to the points to approach them to the real experiment. Four models were applied for simulation and up to five models for fitting.

The analysis of the results shows that with the nowadays state-of-the-art measurement and data treatment techniques, in most of the cases it was possible to distinguish more complex and also more probable bi-ligand and mixed metal–ligand complexes from the simpler 1:1 metal–ligand systems. Statistical evidences to validate the right model were given. Its applicability has been confirmed by generating a similar data mining server (DMS) rule.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Voltammetry; Metal–ligand complexing parameters; Multiligand systems; Mono- and bi-ligand complexes; Mixed-ligand(s) complexes; PROSECE; DMS

1. Introduction

As written in numerous papers, there has been a lot of work done and many efforts put in measuring metal–ligand interactions in model solutions and natural waters, while it is evident that the metals speciation is a good indicator of their bioavailability and hence, their toxicity [1–7]. Voltammetric methods with standard addition of metal ions of interest have shown to be one of the most non-destructive and subtle techniques, able to recognize metal–ligand complexation by directly measuring the free and labile fraction of metal and distinguishing free metal and labile from inert metal complexes [8–12]. There are numerous models which deal with metal speciation with natural organic matter that have become sophisticated taking into account more and more components and parameters affecting the metal speciation and conditions when the bioavailable species could be present, as well as the models of natural organic matter description, which is heterogeneous and complex in its nature. They use large database knowledge about thermodynamic stability constants when discrete ligand approach is applied [13–18] or when continuous functions are associated with NOM properties and affinity to metal, major cations or proton [19–23], both needing a verification in experiments of various techniques. By the combination of such
model—experiment approach, one can advance in characteriza-
tion of samples of interest and obtain a more global speciation conception. In concrete applications, i.e. experi-
mental verifications, the complexity of the system imposed the need for simplification [24,25], a stepwise approach. In-
deed, it has been shown that many situations and experi-
mental data could be explained and covered by a simpli-
fied conception of 1:1 metal–ligand complexes with their
numerical indications of capacity and conditional complex
stability constants [26–33]; on the other hand for a more
complex conception there were no means of enough sensi-
tive measurements and calculations, regarding voltammetric
method approach [34,35]. When facing the complex struc-
ture of the real sample matrix, we always have to bear in
mind that from single technique measurements (voltammet-
ric or any other) one can get information about the so called
conditional stability constants, which does not diminish the
significance of the analyses, but necessitates further interpre-
tation.

Ligands with strong complexing abilities will tend towards
making bi-ligand complexes with trace metals, especially in
the situations of low concentrations of the metals available
[18,24,34,36–46]. In natural water systems, which contain
different ligands, mixed metal–ligand complexes are more
probable to occur than the pure single metal–ligand ones [47].
Their quantitative characterization is not an easy task and is
not often found in the literature. In contrast, theoretical calcu-
lations based on the knowledge of simpler systems, analyses
of factors affecting the recognition of ternary complexes, and
the reviews about particular ligand’s behaviour in the envi-
rionment (e.g. [48–52]) can be found. Mixed metal–ligand
complexes were studied in order to improve the sensitivity
of the voltammeter as electroanalytical method by multiple
enhancement of metal–mixed ligands complexes adsorption
onto mercury drop electrode and formation of metal–mixed
ligands–surface complex, known as synergetic adsorption
[53–55]. Also, this could be a possible type of process (among
these well known) for metals removing from aqueous sys-
tems. Experimental results on synergetic adsorption phenom-
ena of uranium and copper at the hanging mercury drop
electrode have been published [53–55]. If we broaden the
problematic to the formation of metal–ligand–surface com-
plexes (surfaces found in natural waters and sediments and
surface adsorptions acting as a ligand), there is a growing
motivation for the study of ternary complexes [56].

In our previous paper, a detailed analysis of the state-of-
the-art in metal complexing parameters determination in nat-
ural water systems for 1:1 metal–ligand systems has been
given [57]. Starting from these results and using the op-
timal conditions described, possible to perform with our
nowadays voltammetric experimental setup [58] (i.e. non-
linear, logarithmic standard additions) we wanted to analyze
in more detail the ability of our experimental conditions
and analyzing tools to distinguish the situations different
from 1:1 metal–ligand situations that might occur in nat-
ural water systems, such as one and two 1:2 metal–ligand
systems, as well as mixed ligand systems with the studied
metal.

2. Simulation of experiment and formation of data
sets for fitting

The simulation of experiment, i.e. numerical generation of
data points that strictly imitate the distribution and the num-
ber of standard additions, that takes into account the existing
sensitivity of the available instruments and noise distribution
curve, is useful in predicting the recognition of information
comprised in real, experimental data points, as shown previ-
ously [57,59]. According to the design of experimental stan-
dard additions, each point representing the sum of the free
and the labile fraction of added metal was calculated by the
program MINEQL [60].

The concentration ranges, the number of additions of
titrated metal as well as the logarithmic type of additions
were taken from the previous results [57], i.e. for one and
two ligand cases which were analyzed in this work, the con-
centration range of added metal from 1 to 457 nmol L
−1
is covered by 25 standard addition points with constant incre-
ments of pM
1, i.e.

\[ M_{T,i} = 10^{-pM_{1,i}+(pM_{1,i}−pM_{1,1})×i/n} \]

where \( n \) is 25 and \( M_{T,0}, M_{T,25}, \) and \( M_{T,50} \) are the initial, the final
and \( r \)th total metal concentrations, respectively.

This design of titration, named logarithmic, has been the
subject of the development of an automatic apparatus, with
DPASV measurements of labile trace metal concentrations
[61].

Considering the non-uniform repartition of the standard
development of the measurements obtained on real logarithmic
addition titration of cadmium [61], it seemed interesting to try
reproducing that type of variation and apply it in this study.
The noise added to the simulated data before fitting the best
corresponds to:

\[ \epsilon(M) = \frac{1}{\exp((M/M_0) × 10^(-1)) - (ε_0 - ε_{\infty})/((ε_0 - ε_{\infty}))} \]

where \( M \) and \( M_0 \) are the concentrations of metal added dur-
ing and at the end of the titration, respectively; \( ε_0 \) and \( ε_{\infty} \) the
noise limits at the beginning and end of the titration, 20 and
1%, respectively. So the added noise depends on metal addi-
tion and consists of adding random noise to the data points
between \( −\epsilon(M) \) and \( \epsilon(M) \). This process of data modification
by noise addition is presented in Fig. 1.

Taking into account that analytical techniques used for the
characterization of natural ligand complexing properties, es-
pecially when real samples are concerned, can seldom detect
the concentration values lower than 0.1 nmol L
−1
, this value
was put as a limit to remove all the data points obtained by cal-
culation which are lower than 0.1 nmol L
−1
, i.e. the detection
limit (DL) = 1 × 10⁻¹⁰ mol L⁻¹. So, the number of points of one data set exposed to fitting was not always 25; it came down even to 8 for the experiments where the complexation intensity was the strongest. In Fig. 1, 14 points are left after noise application and detection limit restriction.

The existence of two ligands (L) was simulated, a strong (X) and a weak (Y) of 30 and 300 nmol L⁻¹ concentrations, respectively. The position of the points selected for simulation in log(β₁) × log(β₂) space, representing the first and the second stability constants between metal (M) and ligand (L = X or Y) (β₁ = [ML]/[M][L]; β₂ = [ML₂]/[M][L]²), is presented in Fig. 2. These values have been chosen taking into account the results obtained in the literature (or on real systems) [27–29, 31–33, 57, 39–41, 62]. The stability constant values chosen for simulation are summarised in Table 1.

Regarding the formation constants of mixed complexes, i.e. MXY, it has been shown [50, 51] that they should be related to the formation constants of the bi-ligand complexes by a relation of the following kind:

\[ \beta_{XY} = \omega \times (\beta_{2X} \times \beta_{2Y})^{1/2}, \]

with \( \omega \geq 2 \). In our analysis, we used three different values of \( \omega \) for the same combinations of formation constants as those used for simulation of single complexes, i.e. 2, 2 × 10⁻³, and 2 × 10⁻⁶ (Table 1).

Four different groups of models were simulated using MINEQL: S₁ = ML (for X and Y); S₂ = ML + ML₂ (for X and Y); S₄ = MX + MX₂ + MY; and S₅ = MX + MX₂ + MY + MY₂ + MXY. The preset values of complexing parameters used for simulation can be read out from Fig. 2 and from Tables 2–5, respectively. For example: data set 7 stands for log(β₁) = 10, log(β₂) = 16, L = X = 30 nmol L⁻¹, the relevant results are found in Table 3.

For each group of simulated points, three forms of data sets were prepared, namely: (a) for all 25 originally calculated data points, (b) for data points after removal of the points lower than the selected determination limit DL, in our case DL = 1 × 10⁻¹⁰ M (number of data points diminished up to 8), and (c) for data points obtained after removal of data <DL and application of the random noise.

3. Fitting strategy

Also resulting from our previous paper [57], the program PROSECE and its complete procedure was applied in all the fittings. Besides its proven suitability shown in the previous analyses, its structure is much more flexible and able to adapt itself to various fitting models used in this study compared to the other programs that could have been applied instead [59]. Its drawback is, however, the lack of fitted parameters confidence limits estimate.

In order to limit the total number of fittings, which in such analyses easily becomes too big and the obtained results not easy to elaborate, the schedule of fitting generally comprised the models similar or simpler than those used to simulate the data sets. Altogether five different fitting models were applied to data sets, namely, F₁ = ML; F₂ = ML + ML₂; F₃ = MX + MY; F₄ = MX + MX₂ + MY and F₅ = MX + MX₂ + MY + MY₂ + MXY.

In the case of the “right” (matching) fitting model with the simulation model, one can discuss the errors of fitting, i.e. the differences between the preset and the obtained metal–ligand complexing parameters, but since in general we do not know which model is the right one, we should pay more attention to the other parameters of fitting, such as the amount and the shape of the residual of fitting. The parameter that also seemed to be interesting and drawing attention when analyzing the residuals was the number of zero crossings of the residual.

![Fig. 1. An example of modification of data simulated by MINEQL after noise addition and removal of the points below 1 × 10⁻¹⁰ mol L⁻¹.](image1)

![Fig. 2. The legend of distribution of metal first and second stability constants with stronger (X = 30 nmol L⁻¹) and weaker (Y = 100 nmol L⁻¹) ligand used to simulate the experiments.](image2)

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Values of stability constants of mono-, bi- and mixed complexes used for simulation in various combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(β₁)</td>
<td>log(β₂)</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>24</td>
</tr>
</tbody>
</table>
Table 2
Results obtained with PROSECE fitting of the six S1-type experiments using F1, F2 and F3 complexing models on c-type of data sets (n: number of simulated points after removal of points >0.5; X1: legend concentrations in mol L⁻¹; res, normalized residuals of PROSECE fitting)

<table>
<thead>
<tr>
<th>S1 sim</th>
<th>Preset values</th>
<th>F1 = ML</th>
<th>F2 = ML+ML₂</th>
<th>F3 = MX+MY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log(β₁)</td>
<td>n</td>
<td>log(β₂₁)</td>
<td>l₁</td>
</tr>
<tr>
<td>X</td>
<td>8</td>
<td>30</td>
<td>25</td>
<td>8.0</td>
</tr>
<tr>
<td>Y</td>
<td>a</td>
<td>6</td>
<td>300</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>7</td>
<td>300</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>8</td>
<td>300</td>
<td>8.0</td>
</tr>
</tbody>
</table>

Table 3
Results obtained with PROSECE fitting of the 18 S2-type experiments using F1, F2 and F3 complexing models on c-type of data sets; other parameters are the same as in Table 2

<table>
<thead>
<tr>
<th>S2 sim</th>
<th>Preset values</th>
<th>F1 = ML</th>
<th>F2 = ML+ML₂</th>
<th>F3 = MX+MY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log(β₁)</td>
<td>n</td>
<td>log(β₂₁)</td>
<td>l₁</td>
</tr>
<tr>
<td>a + 6</td>
<td>8</td>
<td>16</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>a + 7</td>
<td>8</td>
<td>24</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>a + 8</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>a + 10</td>
<td>12</td>
<td>20</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>a + 12</td>
<td>12</td>
<td>24</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>b + 6</td>
<td>8</td>
<td>16</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>b + 7</td>
<td>8</td>
<td>24</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>b + 8</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>b + 10</td>
<td>12</td>
<td>20</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>c + 6</td>
<td>8</td>
<td>16</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>c + 7</td>
<td>8</td>
<td>24</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>c + 8</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>c + 10</td>
<td>12</td>
<td>24</td>
<td>30</td>
<td>38</td>
</tr>
</tbody>
</table>
4. Results and discussion

Before we start the fitting procedure that should comprise several models, we should try to globally estimate the results by watching the shape of the data expressed in terms of several models, we should try to globally estimate the residuals by watching the shape of the data expressed in terms of the fitting models. F1 = ML, F2 = ML + ML^2, F3 = MX + MY, F4 = MX + MY and F5 = MX + MY + MY + MX; other parameters are the same as in Table 2.

<table>
<thead>
<tr>
<th>S5 sim.</th>
<th>n</th>
<th>F1, res,</th>
<th>F2, res,</th>
<th>F3, res,</th>
<th>F4, res,</th>
<th>F5, res,</th>
</tr>
</thead>
<tbody>
<tr>
<td>d = 4</td>
<td>2</td>
<td>23</td>
<td>4.356</td>
<td>2.708</td>
<td>1.100</td>
<td>0.930</td>
</tr>
<tr>
<td>d = 4</td>
<td>2 × 10^3</td>
<td>12</td>
<td>4.501</td>
<td>0.247</td>
<td>0.406</td>
<td>0.824</td>
</tr>
<tr>
<td>d = 4</td>
<td>2 × 10^6</td>
<td>11</td>
<td>3.646</td>
<td>3.422</td>
<td>0.211</td>
<td>0.191</td>
</tr>
<tr>
<td>f = 2</td>
<td>8</td>
<td>2.984</td>
<td>2.984</td>
<td>0.694</td>
<td>1.612</td>
<td>0.576</td>
</tr>
<tr>
<td>f = 6</td>
<td>2 × 10^3</td>
<td>8</td>
<td>3.547</td>
<td>2.411</td>
<td>2.732</td>
<td>0.795</td>
</tr>
<tr>
<td>f = 6</td>
<td>2 × 10^6</td>
<td>8</td>
<td>2.656</td>
<td>2.656</td>
<td>1.393</td>
<td>0.595</td>
</tr>
<tr>
<td>h = 4</td>
<td>2</td>
<td>9.316</td>
<td>9.316</td>
<td>1.085</td>
<td>1.305</td>
<td>1.001</td>
</tr>
<tr>
<td>h = 4</td>
<td>2 × 10^3</td>
<td>11</td>
<td>6.013</td>
<td>6.013</td>
<td>1.008</td>
<td>3.567</td>
</tr>
<tr>
<td>h = 4</td>
<td>2 × 10^6</td>
<td>11</td>
<td>6.504</td>
<td>6.504</td>
<td>3.604</td>
<td>1.344</td>
</tr>
<tr>
<td>j = 10</td>
<td>2</td>
<td>7.084</td>
<td>7.084</td>
<td>6.883</td>
<td>0.545</td>
<td>0.560</td>
</tr>
<tr>
<td>j = 10</td>
<td>2 × 10^3</td>
<td>11</td>
<td>7.015</td>
<td>7.015</td>
<td>6.834</td>
<td>0.748</td>
</tr>
<tr>
<td>j = 10</td>
<td>2 × 10^6</td>
<td>11</td>
<td>7.037</td>
<td>7.037</td>
<td>6.709</td>
<td>0.472</td>
</tr>
<tr>
<td>i = 12</td>
<td>2</td>
<td>8.021</td>
<td>7.900</td>
<td>1.957</td>
<td>8.088</td>
<td>1.969</td>
</tr>
<tr>
<td>i = 12</td>
<td>2 × 10^3</td>
<td>8</td>
<td>6.697</td>
<td>6.697</td>
<td>5.035</td>
<td>5.609</td>
</tr>
<tr>
<td>i = 12</td>
<td>2 × 10^6</td>
<td>8</td>
<td>7.051</td>
<td>7.051</td>
<td>5.634</td>
<td>6.879</td>
</tr>
</tbody>
</table>

The normalized residuals of all fittings for the data sets (c) (data points obtained after removal of data <DL and application of the random noise) of the four groups of simulated S1, S2, S4 and S5 models fitted by up to five F1–F5 fitting models, are shown in Fig. 3 and will be discussed separately. In general, at first sight, for the S1 and S2 cases the residuals are similar for all fittings, in which case, with the help of statistics we decide whether to accept a more complicated model or stay to the simpler one. For the S4 and S5 cases, we can notice that the application of the right model can in average be distinguished from the wrong ones by comparing the normalized residual values. In the tables the matching model is marked with an asterisk (*).

4.1. One-ligand systems (S1 and S2 simulations)

One-ligand system was presented with both ligands, L = X and Y, and their corresponding stability constants for monoligand as well as bi-ligand cases. Each of the simulated data sets was fitted by three different models, F1 = ML, F2 = ML + ML2 and F3 = MX + MY.

4.1.1. S1 simulation

Six different data sets were considered, three for stronger (X) and three for weaker (Y) ligands; the preset values of complexing parameters, capacity and stability constants are given in Table 2 and can also be read out from Fig. 2. The results of (a) and (b) fittings with ML model were rather

C. Garnier et al. / Analytica Chimica Acta 538 (2005) 263–271
Fig. 3. Residuals of fitting for data sets obtained after removal of data below the DL and application of random noise. S for simulation and F for fitting models: S1 = F1 = ML; S2 = F2 = ML + ML; S3 = F3 = MX + MY; S4 = F4 = MX + MX + MY; S5 = F5 = MX + MX + MY + MY + MY. Legends correspond to tables notation and are named from left to right for each F: S1: 1, 2, 3, a, b and c; S2: 4, 5, 6, 7, 8, 9, 10, 11, 12, d, e, f, g, h, i, j, k and l; S4: a + 4, a + 6, a + 8, a + 10, a + 12, b + 4, b + 6, b + 8, b + 10, b + 12, c + 4, c + 6, c + 8, c + 10, c + 12, d + 4, d + 6, d + 8, d + 10, d + 12, e + 4, e + 6, e + 8, e + 10, e + 12, f + 4, f + 6, f + 8, f + 10, f + 12, g + 4, g + 6, g + 8, g + 10, g + 12, h + 4, h + 6, h + 8, h + 10, h + 12, i + 4, i + 6, i + 8, i + 10, i + 12, j + 4, j + 6, j + 8, j + 10, j + 12, k + 4, k + 6, k + 8, k + 10, k + 12, l + 4, l + 6, l + 8, l + 10, l + 12.

straightforward, almost perfectly matching the preset values, and are not surprising and to be discussed. The results of the simulated points fitting after removal of <DL points and noise insertion with three different models are shown in Table 2. Concerning normalized residuals (residuals divided by number of points), all the three fittings have similar residuals. In that case the simplest model should be accepted, which is the right decision according to the F-test [68]. It is interesting that for F2 model, the second constant $\beta_2$ has just been figured out by the fitting program, with almost no influence on the result of total ligand (except for the Sim. a case). The F3 model applied on strong ligands gives unacceptable results, but when applied on weaker ligands it just divides the total ligand in two groups of almost equal constants, which points that the decision of accepting the F1 model as the right one would be the correct choice.

4.1.2. S2 simulation

Eighteen different data sets were simulated, 9 for stronger (X) and 9 for weaker (Y) ligands. The results of data fitting
from the (c) cluster with the same three models as in the
S1 case are shown in Table 3. Looking only to the residu-
als provides no enough evidence to determine which model
for fitting is the right one. On the contrary, for the set of
stronger (X) ligands, lower residuals were obtained with a
model offering two, rather than one ligand with two constants
for mono- and bi-ligand complexes. In the group of S2 sim-
ulations, there is expressed the strongest impact of passing
from fitting of all 25 data points (a), to fitting of data points
without those eliminated by the DL (b) and fitting of noisy
data (c). By analyzing only the (c) cluster, one can hardly
decide about the right model and consequently about the val-
ues of the complexing parameters. In such a case the answer
lies in redesigning the experiments either by measuring more
replicates or reﬁning and increasing the number of standard
additions in order to increase the detection limit and lower the
noise.

4.2. Two-ligand systems (S4 and S5 simulations)

Two-ligand systems were presented with the mixtures of
ligands X and Y with the metal M and stability con-
stants for mono- bi-ligand cases as well as mixed
metal–ligand complexes. Each of the simulated data sets
was ﬁtted by four or ﬁve various models, \( F_1 = M \), \( F_2 = M + M \), \( F_3 = M + Y \), \( F_4 = M + M + M + M \)
and \( F_5 = M + M + Y + M + M + Y \).

4.2.1. S4 simulation

Fifteen different data sets were simulated following the
S4 model, i.e. there is one stronger (X) ligand that forms
mono- and bi-ligand–metal complexes and one weaker (Y)
ligand that forms mono complex. According to their simula-
tion names given in Table 4 and Fig. 2, one can read out their
ligand concentrations and stability constants used. In Table 4,
the obtained ﬁtting parameters for the matching model as well
as the values of the residuals for all the models applied are
given together with the number of points of each data set for
the (c) cluster.

In 9 out of the 15 cases the normalized residual was the
smallest for the matching model. The weaker ligand concen-
tration was retrieved with better score than the stronger one,
with less than 10% of error, 11 out of the 15 and 3 out of the
15, respectively. Stability constant of the weaker ligand was
well retrieved in all the 15 cases, while for the stronger ligand
this happened in only 3 out of the 15 cases. Those positive
cases were characterized at the same time with the number of
points considerably high, i.e. the constants were the lowest
of the proposed constant sets.

4.2.2. S5 simulation

Fifteen different data sets were simulated following the
S5 model, i.e. there is one stronger ligand that forms mono-
and bi-ligand–metal complexes, one weaker ligand that forms
mono- and bi-ligand–metal complexes, and there is one
mixed-ligand complex. According to their simulation names
given in Table 5, Fig. 2 and Table 1, one can read out their
ligand concentrations and stability constants used. In Table 5
the obtained ﬁtting parameters for the matching model as well
as the values of the residuals for all the models ap-
plied are given together with the number of points of each data set for
the (c) cluster. In 10 out of the 15 cases the nor-
malized residual was the smallest for the matching model,
which is a better score than in the case of S4 simulation.
However, the values of complexing parameters were esti-
mated with higher tolerances. Ligand concentrations were
underestimated, weaker ligands in higher extent than the
stronger ones, in more than half of the cases. From ﬁve of
the simulated stability constants the best retrieved constants
after noising procedure were \( \beta_{1X} \) and \( \beta_{2Y} \), the constants
for weaker mono- and bi-ligand–metal complexes. It is evident
that for such rather complicated model, broader range of val-
ues of simulated parameters should be selected and the sim-
ulated data consequently treated in order to survey a speci-
ciﬁc rule and/or a range of more certain recognition of preset
parameters.

Starting from these results, two characteristic cases were
studied in more detail, i.e. the case \( d + 4 \), \( \omega = 2 \) and the case \( f + 6 \), \( \omega = 2 \times 10^6 \). Three-fold repetition of the experiments
was simulated in order to test its inﬂuence on the correct
retrieving of the parameters. The results given in Table 6
brought to the conclusion that measurement repeating in great
extent eliminates the inﬂuence of the noise (seen from the
comparison of the corresponding (b) and (c) set of results),
but cannot compensate for the missing of the data points \( \Delta DL \).

\( \Delta DL \) comparison of a-type of data with b- and c-types in Table 6a,
which in fact carry the major part of information about the
strong complex parameter values.

4.3. Data mining server rules

Handling with a big quantity of calculated data, we tried
to possibly ﬁnd a general rule that could help us recognize
the right model of ﬁtting, in our case simulated data, but in
general, measured data of some unknown samples.

To that purpose we have prepared our data according to
the rules of the data mining server (DMS) [69], where data
analysis is performed based on knowledge induction by the
inductive learning by logic minimization (ILLM) system.
Using known data sets and their classiﬁcation this program
searches for classiﬁcation rules that could be applied to the
new data sets of the same kind. An attempt of mixing (a),
(b) and (c) sets of data did not give any meaningful result,
so we restricted ourselves to the (c) set. A table with 239 ex-
ample rows (containing all the mentioned sets of simulated
data ﬁtted to all the mentioned models) was constructed, 70
of them was found in the positive class, which in our case was
a matching simulation and ﬁtting model. Deﬁned attributes
were all preset parameter values of metal complexing capaci-
ty and their mutual differences and products, the number of
data points in a particular set, the results of ﬁtting including
the amount of residual of ﬁtting, and the residual normalized
Sim. Fitting model $n$ Type $n$ of meas. $\text{res}_n$ $\log(f_{X}(X))$ $\log(f_{Y}(X))$ $\log(f_{Y}(Y))$ $\log(f_{XY}(X, Y))$ $X_1$ $Y_1$

$F^5 + 4n = 2$

$F^5$ 25 a 1 0.194 8.7 16.1 6.0 12.7 14.4 17.4 27
$F^5$ 23 b 1 0.056 8.4 16.1 6.0 12.5 14.6 21.6 286
$F^5$ 23 c 1 0.908 8.5 16.1 5.8 12.1 12.2 23.9 368
$F^5$ 69 b 3 1.066 8.4 16.1 5.9 12.2 14.2 24.0 341
$F^5$ 23 (av) c 3 0.623 8.5 16.1 5.7 12.2 8.4 23.9 383

$F^5 + 6n = 2 \times 10^6$

$F^5$ 25 a 1 3.071 3.6 25.2 5.4 16.1 26.3 31.2 284
$F^5$ 8 b 1 0.583 6.7 16.6 9.6 16.7 24.2 17.5 156
$F^5$ 8 c 1 0.765 5.9 18.7 9.6 16.5 18.3 16.8 152
$F^5$ 24 c 3 1.326 7.1 15.9 9.6 16.7 20.9 19.9 157
$F^5$ 8 (av) c 3 0.663 7.2 16.1 9.6 16.7 21.6 19.2 157

This study showed that by means of voltammetric instrumentation supplied with automatic burettes and sophisticated data treatment that enables quick scanning through different fitting models, it is possible to conclude more precisely about the metal–ligand structure in a measured sample. However, for more accurate values of the complexing parameters, a more detailed study should be undertaken including repeating of experiments, and/or putting denser metal additions (in cases of strong ligands when larger number of first additions are so complexed to leave the free metal below the detection limit), considering that rather a subtle structure of dissolved metal–ligand complexes is tried to be distinguished.

Analyses by the method of data mining server confirmed the statistical methods of residuals analysis and have shown to be applicable in such simulation—fitting modelling systems as a prediction tool and help as one of the criteria in data interpretation. A more specific rule is to be expected with a denser web of modelled parameters, which could be the aim of some future work.

**Acknowledgement**

The authors wish to thank the Ministry of Science, Education and Sport of the Republic of Croatia and the Ministry of National Education, Teaching and Research of France. Also we thank the Ph.D. school of the University of Sud Toulon Var for financial support of collaboration between France and Croatia and D. Gamberger for his help in explanation and application of DMS rules. We would like to thank two anonymous reviewers for their comments and suggestions.

**References**
