SUPPORTING INFORMATION

On Modeling Complexometric Titrations of Natural Water Samples

R.J.M. Hudson, E.L. Rue, and K.W. Bruland

Table of Contents

1) Raw Data from Cu Titrations of Natural Water Samples

A) Kozelka and Bruland ASV analysis. The ASV analysis was performed by Kozelka in a field lab on a fresh, filtered sample from the "mid-bay" site in June 1994. The seawater had a salinity of 28.6, pH of 7.8, and ambient dissolved $[Cu]_T$ = 16.1 nM (*1*). Details of the analytical methods are presented in (*1*). Since no concurrent determination of S^{REF} was available, it was estimated from S^{SIC} and the ratio S^{SIC}: S^{REF} = 0.73 observed in subsequent analyses of frozen samples from a nearby site made using the same instrument. This yielded S^{REF} ~175 nA/nM.

The following data were previously published in graphical form as Fig. 2 of Kozelka and Bruland (*1*).

Table S-1. Original ASV Data provided by P. Kozelka.

[Cu] Added (nM)	Peak Current (nA)
0	0
10	150
20	500
30	1280
40.4	1420
60	2770
75.6	4320
100.7	6500
120	8260
150	12060
185	16500

B) Rue and Bruland CSV Analysis: CLE-ACSV complexometric titrations were performed by Rue at UCSC on a filtered and frozen sample from the lower Narragansett Bay study site (*2*). The original sample had a salinity of 30.3, original pH of 7.9, and ambient dissolved $[Cu]_T$ of 12.7 nM (*3*). The sample pH was buffered at 8.0 using 6.5 mM HEPPS and 5 titrations were conducted at added ligand (salicylaldoxime, SA) concentrations of 1.1, 2.75, 5.5, 11, and 55.5 μM. Details of the analytical methods are presented in (*2*). S^{REF} determined concurrently in UVSW was 4.4 nA/nM/min. Peak heights were linear from 0.5-1 minutes of deposition at -0.05 V.

Calculated concentrations of $Cu(SA)_2$ from two of the following titrations were previously published in graphical form (Fig. 2A of ref 2).

Table S-2. Original CSV data provided by E.L. Rue.

[SA] (μM)	[Cu] Added (nM)	Time (min)	Original log α′	Original S	Peak Current (nA)
1.1	0	0.5	3.1	0.83	0
1.1	1.8	0.5	3.1	0.83	0
1.1	3.6	0.5	3.1	0.83	0.5
1.1	5.4	0.5	3.1	0.83	1
1.1	9	0.5	3.1	0.83	3
1.1	20	0.5	3.1	0.83	6
1.1	40	0.5	3.1	0.83	18
1.1	60	0.5	3.1	0.83	35
1.1	80	0.5	3.1	0.83	51
2.75	0	1	3.9	2.5	0
2.75	1.8	1	3.9	2.5	0
2.75	3.6	1	3.9	2.5	2
2.75	5.4	1	3.9	2.5	6
2.75	7.2	1	3.9	2.5	12
2.75	9	1	3.9	2.5	15
11	0	1	5.1	2.5	5
11	1.8	1	5.1	2.5	7
11	3.6	1	5.1	2.5	10
11	5.4	1	5.1	2.5	16
11	9	1	5.1	2.5	23
11	20	1	5.1	2.5	48
11	40	1	5.1	2.5	99
27.5	0	1	5.9	2.5	10
27.5	5.4	1	5.9	2.5	22.5
27.5	9	1	5.9	2.5	35
27.5	20	1	5.9	2.5	60
27.5	40	1	5.9	2.5	110
55.5	0	1	6.5	3.3	23
55.5	1.8	1	6.5	3.3	29
55.5	3.6	1	6.5	3.3	36
55.5	5.4	1	6.5	3.3	43
55.5	20	1	6.5	3.3	91
55.5	40	1	6.5	3.3	155

C) Moffett CSV Analysis: CLE-ACSV complexometric titrations were performed by Moffett on a fresh sub-sample taken from the same large-volume sample as Rue and Bruland above (*2*). Three titrations were conducted with benzoylacetone (bzac) added at concentrations of 100, 250, and 500 (μM). Details of the analytical methods are discussed in (*2*). Analyses of Cu standard additions to a UV-oxidized sample yielded S^{REF} = 0.64-0.78, but currently available information does not indicate whether they correspond to the same instrument settings as used for the titrations.

Calculated concentrations of $Cu(bzac)_2$ from two of the following titrations were previously published as Fig. 2B of (*2*). The entire data set was also modeled in that paper. The original calculations used S^{ICI} = 0.4 for all analytical windows, a value was presumably derived from internal calibration of the titration at $[AL]_T$ = 500 μM.

Table S-3. Original CSV Data provided by J.W. Moffett.

[bzac] (μM)	Dep. Time (min)	[Cu] Added (nM)	Peak height [a]
100	1	10	0.25
100	1	20	2.5
100	1	25	3
100	1	30	5.5
100	1	40	7.75
100	1	50	14
100	1	75	19
100	1	100	26
250	1	0	0.25
250	1	2	0.38
250	1	4	0.38
250	1	10	1.5
250	1	20	4
250	1	25	7
500	1	0	0.75
500	1	4	1.5
500	1	6	3
500	1	10	4
500	1	15	6.5

[a] Peak heights reflect measurements from chart recorder output in units of length, e.g, cm. Values are proportional to current.

D) Campos and Van den Berg CSV Analysis:

CLE-ACSV complexometric titrations were performed by these investigators on a sample collected from the NW Mediterranean Sea. The seawater had a salinity of ~38 and was buffered to pH 8.35 using 10 mM borate. The ambient dissolved $[Cu]_T$ equaled 3.1 nM. Details of the analytical methods are presented in *(5)*. Since no concurrent determination of S^{REF} is available, it was estimated as 1.5 nA/nM/min from measurements of S^{REF} on other dates as reported in Fig 3B of *(5)*.

Table S-4. Original CSV Data provided by M.L.A.M. Campos. Deposition time = 1 min, potential = -0.15 V *(5)*.

[SA] (μM)	[oxine] (μM)	[Cu] Added (nM)	Original S	Peak Current [a]
1	0	7.1	0.426	0.095[b]
1	0	10.6	0.426	0.48
1	0	13.1	0.426	0.75
1	0	15.6	0.426	1.45
1	0	18.1	0.426	1.719
1	0	23.1	0.426	3.94
1	0	28.1	0.426	6.023
1	0	33.1	0.426	8.128
2	0	3.1	0.629	0.013[b]
2	0	5.1	0.629	0.314
2	0	7.1	0.629	0.547
2	0	9.1	0.629	1.021
2	0	10.6	0.629	1.448
2	0	13.1	0.629	2.24
2	0	15.6	0.629	2.68
2	0	18.1	0.629	3.781
2	0	20.6	0.629	6.44
2	0	23.1	0.629	7.642
10	0	3.1	0.87	0.43
10	0	4.1	0.87	0.789
10	0	5.1	0.87	2.045
10	0	7.1	0.87	3.082
10	0	9.1	0.87	4.595
10	0	10.6	0.87	6.706
10	0	13.1	0.87	7.588
10	0	15.6	0.87	10.67
0	1	5.1	1.0	3.47
0	1	7.1	1.0	4.58
0	1	9.1	1.0	6.23
0	1	11.1	1.0	8.52
0	1	13.1	1.0	10.47
0	1	15.6	1.0	12.46

[a] Peak currents are reported in units of nA for SA data. Oxine data are normalized by S^{SIC} and therefore have units of nM.

[b] Below reported detection limit of 0.1 nA.

2) Dependence of Sensitivity in Reference Media on $[SA]_T$.

As discussed elsewhere (*4*), the sensitivity in reference media is a function of added ligand concentrations. To account for this effect in multi-window and overload analyses, we used S^{REF} determined at various SA concentrations.

Table S-4. CSV Response at $[Cu]_T$ = 18 nM and various $[SA]_T$ in UV-oxidized seawater at pH 8.0. Measurements by E.L. Rue.

	Peak Current (nA)		
$[SA]_T$ (μM)	**Replicate 1**	**Replicate 2**	**R_{SA}**
1.1	55	56	0.53
2.75	84	86	0.81
11.0	96	102	0.94
27.5	98	100	0.94
55.0	104	106	1.00

Table S-5. Ratio of S^{REF} at $[SA]_T$ to S^{REF} at $[SA]_{T0}$ = 25 μM (R_{SA}) (*4*).

$[SA]_T$ (μM)	**R_{SA}**
1	0.55
3	0.78
5	0.92
25	1.00

Using inverse-square weighting and assuming that values in Table S-5 are the mean of 3 measurements, we fitted the following function to both SA data sets jointly (r^2 = 0.997):

$$S = S_{max} \cdot K \cdot [SA]_T / (1 + K \cdot [SA]_T) \quad \text{(S-1)}$$

S_{max} = 106.1 nA for Rue and Bruland's data; 1.05 for Kogut and Voelker's data

$K = 1.08 \times 10^6\ M^{-1}$

The related function for R_{SA} at $[SA]_T$ in μM relative to an arbitrary $[SA]_{T0}$ is then:

$$R_{SA} = [SA]_T / [SA]_{T0} \times (1+1.08\times[SA]_{T0}) / (1+1.08\times[SA]_T) \quad \text{(S-2)}$$

Fig. S-1 shows the fit of this function to the data.

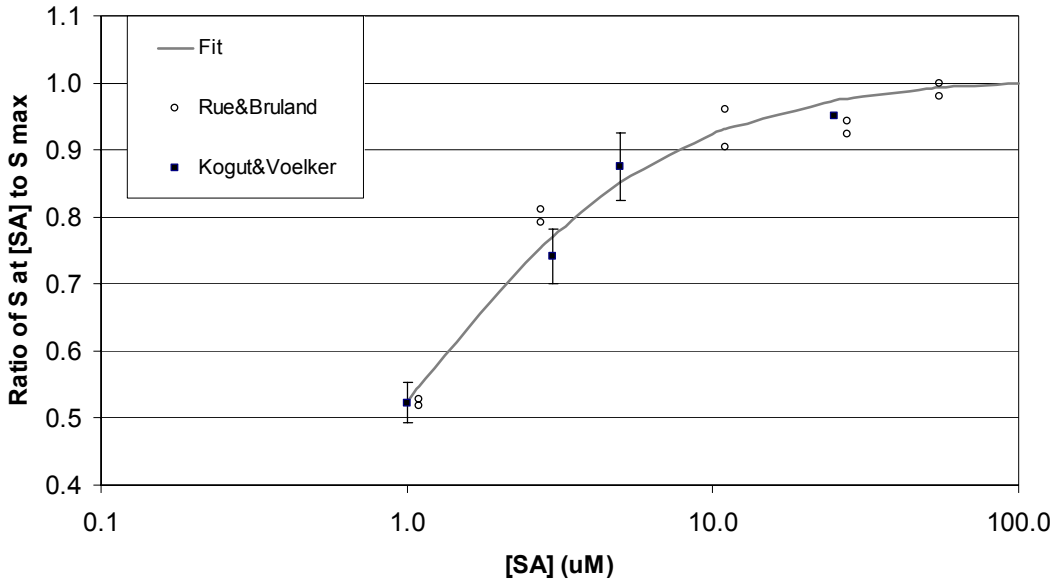

Fig. S-1. Dependence of S^{REF} on $[SA]_T$.

The fact that both data sets fit the same function despite the differences in deposition potential, -0.05V Rue *(2)* versus -0.08 V Kogut *(4)*, is consistent with the small dependence of the sensitivity on deposition potential over the range -0.10 V to – 0.80 V (Fig. 3 in ref *5*). This also justifies applying these results to Campos and van den Berg's multi-window titrations, which used -015 V.

3) Modeling Differences in Surfactant Effects between Analytical Windows

A variety of reasons make it necessary to test hypotheses that affect S in different analytical windows. For example, if titrations with two different ligands are fitted simultaneously, they generally will not have the same S. Or, if one wishes to test the hypothesis that one window does not fit the model well, a parameter that quantifies its divergence could be added to S for that window. Finally, an important assumption made for both the overload approach and the applications of the multi-window models considered thus far is that the dependence of S on $[AL]_T$ (or on the added ligand used) can be defined from R_{AL} measured in the reference medium.

Kogut and Voelker (*4*) have suggested that there may be greater surfactant effects at low $[SA]_T$, which would invalidate this assumption. Actually, this assumption of eq 8 in the paper is not essential for the MG-modeling approach. This is a potentially powerful realization, as it suggests that it is possible to measure the dependence of the surfactant effect (φ) on $[AL]_T$.

An empirical description of the difference of the sensitivity in the j^{th} window (S_j) from the sensitivity (S_0) in the highest α_{MREF} window of the dataset is obtainable by fitting parameters (ΔS_j):

$$S_j = S_0 \cdot R_{AL,j} \cdot \left(1 + w_1 \cdot \Delta S_1 + \cdots + w_{j\max-1} \cdot \Delta S_{j\max-1}\right) \qquad \text{(S-3)}$$

where a dummy independent variable w_j is introduced for each window j with $w_j \equiv 1$ and $w_k \equiv 0$ for all other windows with $k \neq j$. Tests for the statistical significance of any differences in S (ΔS_j) are then possible. A similar approach was used to fit SA and oxine titrations simultaneously.

Not surprisingly, preliminary tests of multi-window models with ΔS_j worked in many cases, but also showed that collinearity between ΔS_j and S_0 can prevent convergence when using eq S-3. In particular, we suggest that when using (S-3), S_0 should be bounded by S^{REF} if known, ΔS_j should bounded by zero or ΔS_{j+1}, and non-significant ΔS_j excluded from the model. Note also that if a mechanistic or even an empirical model defining the relationship between S and $[AL]_T$ can be formulated, it may be possible to reduce the number of parameters and associated collinearity problems. Careful experimental work designed to derive models of the surfactant effect would help immensely. Despite the potential pitfalls of the approach outlined here, it is worth further consideration as there are few other practical approaches to quantifying surfactant effects in natural samples.

4) Example SAS Code for M2 Model

Input File "Narrdata.csv." These are the Rue and Bruland data from above.

```
1,12.7,0,0,1.1
1,12.7,1.8,0,1.1
1,12.7,3.6,0.5,1.1
1,12.7,5.4,1,1.1
1,12.7,9,3,1.1
1,12.7,20,6,1.1
1,12.7,40,18,1.1
1,12.7,60,35,1.1
1,12.7,80,51,1.1
2,12.7,0,0,2.75
2,12.7,1.8,0,2.75
2,12.7,3.6,2,2.75
2,12.7,5.4,6,2.75
2,12.7,7.2,12,2.75
2,12.7,9,15,2.75
3,12.7,0,5,11
3,12.7,1.8,7,11
3,12.7,3.6,10,11
3,12.7,5.4,16,11
3,12.7,9,23,11
3,12.7,20,48,11
3,12.7,40,99,11
4,12.7,0,10,27.5
4,12.7,5.4,22.5,27.5
4,12.7,9,35,27.5
4,12.7,20,60,27.5
4,12.7,40,110,27.5
5,12.7,0,23,55.5
5,12.7,1.8,29,55.5
5,12.7,3.6,36,55.5
5,12.7,5.4,43,55.5
5,12.7,20,91,55.5
5,12.7,40,155,55.5
```

SAS (v. 8.2) Code:

```
Title1 "M2-Model Estimation and Monte Carlo Simulations';

* SAS Code by: Robert J.M. Hudson;
* V. 1.0 December 2002;
* Department of Natural Resources and Environmental Sciences;
* W-503 Turner Hall, 1102 S. Goodwin Ave.
* University of Illinois, Urbana, IL 61801;
* Contact: rjhudson@uiuc.edu;
* Please report problems;

Options ls=72;

data NarragansettSA;
infile 'Narrdata.csv' dlm = ',';

input win Cu0 CuAdd Ip  ALT;

MT = Cu0 + CuAdd;
ALT = ALT/1e6;

if(Ip=0)  then delete;

if(win=1) then win1=1;
else win1=0;
if(win=2) then win2=1;
else win2=0;
if(win=3) then win3=1;
else win3=0;
if(win=4) then win4=1;
else win4=0;
if(win=5) then win5=1;
else win5=0;

factor = 0;
if(win=1)          then RAL = 0.529*0.5;
else if(win=2) then RAL = 0.81;
else if(win=3) then RAL = 0.943;
else if(win=4) then RAL = 0.943;
else if(win=5) then RAL = 1.000;

wt = 1/Ip/Ip;

proc print data = NarragansettSA;
run;
```

```
data NarrMonte;

KMA1 = 9.765;
KMA2 = 15.084;
ALT = 0;
RAL = 1;
wt=1;
do i = 1 to 30;
MT = 4 + i*2;
output;
end;
MT = 12.7;
output;

run;

proc model data = NarragansettSA;
control KMA1 9.567 KMA2 14.572 alphaIN 13;
parms logK1 logK2 logL1 logL2 Ssa LsT aMLw;

* Solve two-ligand problem;
* Code solves for logKi in 1/nM units;
* Code solves for logLi in tenth nM units, i.e., logLi = 2 = 10 nM.;


* Remove comment if Ls is desired;
* MT = max(0,MT - 10**(logLs-1));

* Define ligand parameters;
  K1 = 10**logK1;
  L1 = 10**(logL1-1);
  K2 = 10**logK2;
  L2 = 10**(logL2-1);

* Comment out one of the following;
   alphaMLw = 0;
*  alphaMLw = 10**aMLw;

* Set up calculation in case a significant fraction of added ligand is complexed;
  alphaMAL = 10**KMA1*ALT+10**KMA2*ALT**2;
  alpha = alphaMAL + alphaIN + alphaMLw;

  Aprime    = ALT;
  Aprime0    = max(1e-10,ALT/2);
```

```
* Solve equilibrium speciation model;
  do while (abs((Aprime-Aprime0)/max(Aprime,1e-10)) > 0.0001);
    Aprime0   = Aprime;
    alphaMAL = 10**KMA1*Aprime+10**KMA2*Aprime**2;
    alpha     = alphaMAL + alphaIN + alphaMLw;

    K1=K1/alpha;
    K2=K2/alpha;

* Solve 2-ligand speciation model;

          *   f(M+) = M+^3 + a2 M+^2 + a1 M+ + a0;

          * Positive, real root to cubic equation.;
          * See Gellert, W., Küstner, H., Hellwich, M., Kästner, H.
          * The VNR Concise Encyclopedia of Mathematics;
          * Van Nostrand Reinhold: New York, 1975.
          * See also: http://mathworld.wolfram.com/CubicEquation.html;

    a2 = (L1+L2-MT) + (K1+K2)/K2/K1;
    a1 = (K2*L2+K1*L1-(K1+K2)*MT+1)/K1/K2;
    a0 = -MT/K1/K2;
* Intermediate calculations;
    P = a1/3.-a2**2./9.;
    Q = a1*a2/6.-a0/2.-a2**3/27.;
    D = Q**2+P**3;
    if(D<0) then do;
          theta = arcos(Q/sqrt(-P**3.));
          Mfree = (2.*sqrt(-P)*cos(theta/3.)-a2/3.)/alpha;
    end;
    else do;
          Mfree = ((Q+sqrt(D))**(1/3)+(Q-sqrt(D))**(1/3)-a2/3)/alpha;
    end;
*End 2-ligand speciation model;

end;
    Aprime         = ALT/(1.+alphaMAL*Mfree/1e9);
    alphaMref  = 10**KMA1*Aprime+10**KMA2*Aprime**2 + alphaIN;
    pM = -log10(Mfree)+9;

* End speciation model calculation;

* Calculate observed signal;
    Ip = 10**Ssa*RAL*Mfree*alphaMref;
```

```
 fit Ip start = (Ssa = 0 0.5 1 logK1 = 5 5.5 6 logK2 = 1.5 2 2.5 3 logL1 = 2 2.5 3 logL2 = 3 3.5 4
      4.5 5 logLs = 1 1.5 2 aMLw = 3)/ missing = pairwise outcov outest=ip_est outs=s_data
      converge=0.0001;
 WEIGHT wt;

 solve pM / data=NarrMonte estdata=ip_est random=10000 seed = 14098 out = monte simulate;
 id MT;
run;

proc sort data=monte;
by MT;
run;

proc univariate data=monte;
 by MT;
 var pM;
output out = sumstat median=median q1=q1 q3=q3 pctlpre=P_ pctlpts=2.5,97.5;
 run;

proc print data = sumstat;
run;
```

To solve 1-ligand problems, delete references to K2 and L2 and substitute the following code for the "Solve 2-ligand speciation model" section (p. 11):

```
B   = K1*(L1-MT) + alpha;
Mfree = (-B+(B*B+4*K1*MT*alpha)**0.5)/2/K1/alpha;
```

References:

(1) Kozelka, P. B.; Bruland, K. W. *Mar. Chem.* **1998**, *60*, 267-282.
(2) Bruland, K. W.; Rue, E. L.; Donat, J. R.; Skrabal, S. A.; Moffett, J. W. *Anal. Chim. Acta* **2000**, *405*, 99-113.
(3) Wells, M. L.; Kozelka, P. B.; Bruland, K. W. *Mar. Chem.* **1998**, *62*, 203-217.
(4) Kogut, M. B.; Voelker, B. M. *Environ. Sci. Technol.* **2001**, *35*, 1149-1156.
(5) Campos, M. L. A. M.; Van den Berg, C. M. G. *Anal. Chim. Acta* **1994**, *284*, 481-496.