

Supplementary material

***PEST-ORCHESTRA*, a tool for optimising advanced ion-binding model parameters: derivation of NICA-Donnan model parameters for humic substances reactivity**

Noémie Janot,^{A,B} *José Paulo Pinheiro*,^{A,B} *Wander Gustavo Botero*,^C *Johannes C. L. Meeussen*^D and *Jan E. Groenenberg*^{A,B,E,F}

^ACNRS, LIEC, UMR7360, 15 Avenue du Charmois, Vandoeuvre-lès-Nancy F-54501, France.

^BUniversité de Lorraine, LIEC, UMR7360, 15 Avenue du Charmois, Vandoeuvre-lès-Nancy, F-54501, France.

^CFederal University of Alagoas (UFAL), Campus Arapiraca, 57309-005-, Arapiraca, AL, Brazil.

^DNRG Consultancy & Services, PO Box 25, 1755 ZG Petten, The Netherlands.

^EWageningen University, section soil quality, PO Box 47, 6700 AA Wageningen, The Netherlands.

^FCorresponding author. Email: bertjan.groenenberg@wur.nl

TABLE OF CONTENT

1	Structural design of ORCHESTRA software.....	S3
2	Functional design: Coupling PEST with ORCHESTRA software.....	S6
3	Implementation of the unconstrained procedure in PEST-ORCHESTRA.....	S7
3.1	ORCHESTRA chemistry file (chemitry1.inp)	S7
3.2	ORCHESTRA input file (input.dat)	S9
3.3	PEST chemistry file (chemistry1LFA.tpl).....	S10
3.4	PEST control file (H_LFA.pst)	S11
3.5	PEST instruction file (outputFLFA.ins)	S12
4	Implementation of the master curve procedure in PEST-ORCHESTRA.....	S13
4.1	Description of the master curve approach	S13
4.2	PEST chemistry file (chemistry1.tpl)	S15
4.3	Input file (input1.dat).....	S16
4.4	postprocess.inp	S17
5	Exclusion of electrostatic interactions in the NICA-Donnan model.....	S17
6	Implementation of metal-binding parameters optimization.....	S18
7	Results.....	S19
7.1	Results of proton-binding parameters optimization	S19
7.1.1	Unconstrained procedure.....	S19
7.1.2	Master curve procedure	S20
7.1.3	2-step procedure	S21
7.2	Comparison with generic Cu-binding parameters derived by Milne et al. (2003) ..	S21

1 Structural design of ORCHESTRA software

Within ORCHESTRA three fundamental classes of chemical objects are defined: the entities, the reactions and the phases^[1]. The entities are chemical objects with an activity, a concentration and a mass balance. The reactions define the interaction between different entities, and each is associated with a reaction constant. The phases represent conceptual or physical phases of the system for which the mass balances are calculated.

Within ORCHESTRA the definition of the chemical and transport model equations is separated from the method to solve the equations^[1]. Models are formulated as objects in text format and solved by a generic equation solver (Figure S1), which makes it possible for the user to modify and add new models in a straightforward way. The general equation solver, together with the three object classes, is implemented in a java-executable jar file.

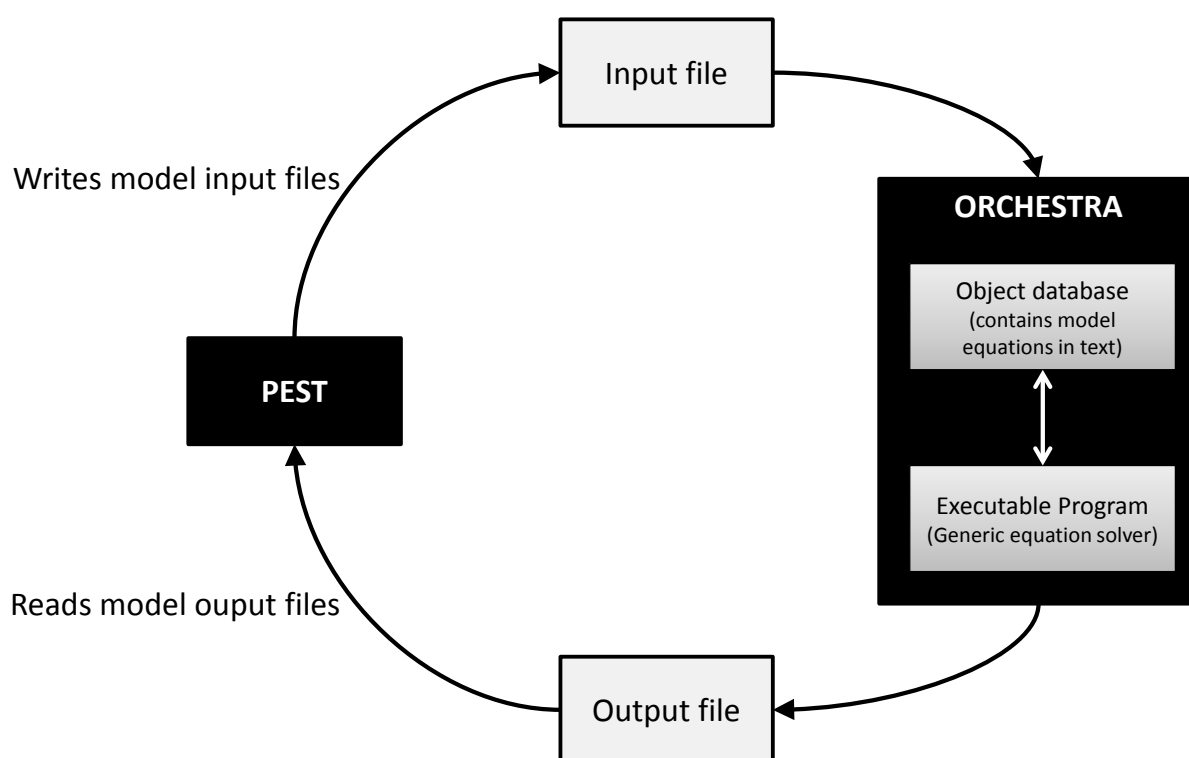


Figure S1. Overview of the structure of the PEST-ORCHESTRA framework. The communication between PEST and ORCHESTRA is organized via the input- and output files of ORCHESTRA.

Within the “composer.inp” file the files making up a particular ORCHESTRA application are defined. The operations and their sequence (i.e. reading from input files, performing calculations and writing to output files) are defined within the concert file “concert.xml”. A typical ORCHESTRA application to perform chemical speciation calculations is built up from (i) a chemistry-file (chemistry1.inp) in which the chemical system is defined in terms of its phase-hierarchy, the entities and reactions; (ii) an input file (input.dat) with the concentrations of the entities (in a certain phase) and (iii) one or more output-files to which model output can be written, *i.e.*, the values of all variables defined in a particular application (e.g. species concentrations, Donnan volume, etc.).

In ORCHESTRA, the phase hierarchy is adjustable by the user; as an example, the phase hierarchy for a humic acid-containing system as used in this study is shown in Figure S2. Within this hierarchy, each phase includes the phase(s) below it, and two consecutive phases are linked by a user-defined factor. Units of the different phases depend on the factor linking the two phases.

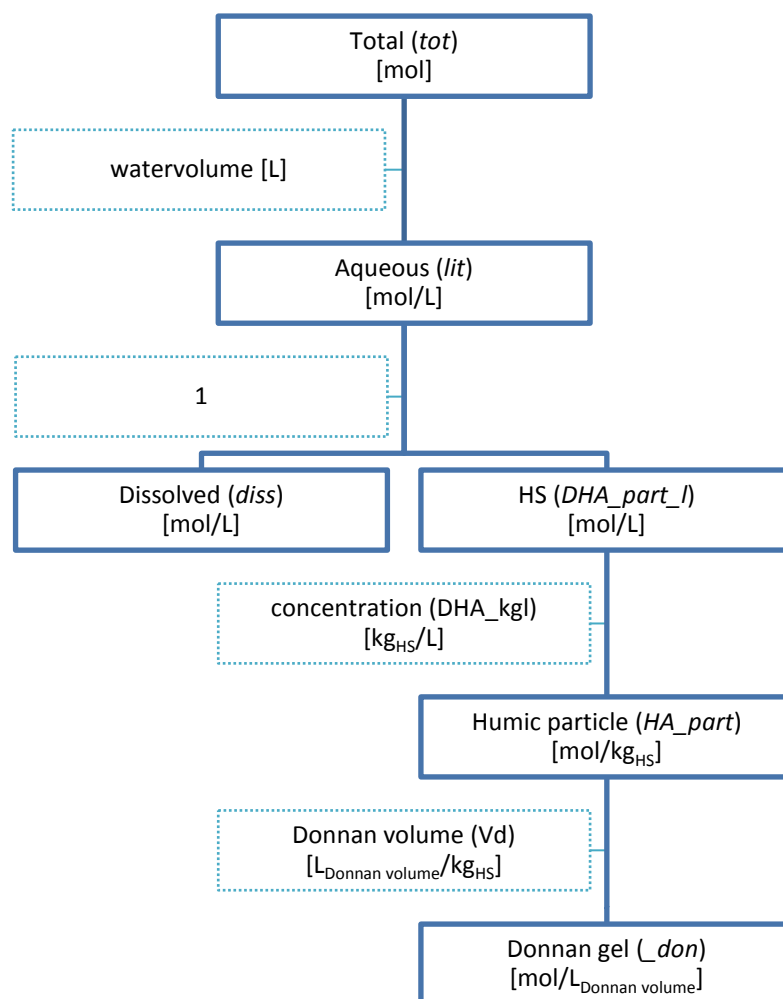


Figure S2. Overview of the phase hierarchy for a humic acid-containing system. In dotted boxes are the factors linking the phases. The names used in ORCHESTRA are written in between parenthesis, the units in between brackets. The phases definition in text format is included in the chemistry1.inp file (see section 3.1)

Several adsorption models have already been defined as object classes (in the file objects2016.txt) for use in ORCHESTRA, including the NICA-Donnan model for humic and fulvic acids and CD-MUSIC for metal oxides^[2]. Generic NICA-Donnan model parameters values for proton and metal binding^[3,4] are provided in the database file adsmodels.txt. Since model types are defined as objects, in text format, it is possible for the user to modify and/or add implementations to the desired model.

2 Functional design: Coupling PEST with ORCHESTRA software

The functionality to optimize variables in ORCHESTRA in order to fit experimental data has been implemented by the coupling of ORCHESTRA with PEST, a model-independent parameter estimation and uncertainty analysis software package. Model parameters are estimated iteratively to obtain the lowest sum of squared residuals using a Levenberg - Marquard algorithm. Here we combined ORCHESTRA version 9 February 2016, downloaded February 2016 from <http://www.meeussen.nl/orchestra/>, with PEST version 13.0, downloaded June 2013 from <http://www.pesthomepage.org/Downloads.php>.

The communication between PEST and ORCHESTRA is organized via the input and output files of ORCHESTRA (Figure S1). Within each iteration cycle PEST writes the input file for ORCHESTRA with therein the estimates of the parameters values to be optimized. The calculations are performed by ORCHESTRA using these parameters values and the results are written to an output file. This output file is read by PEST and the calculations are compared with the observations. The goodness of fit is quantified with the sum of the squared residuals, which is to be minimized. Based on these results PEST gives new estimates for the parameter values. The cycle is repeated until the difference between the parameter values of successive iterations is smaller than a user-defined criterion.

The input file to be read by ORCHESTRA is the chemistry file (chemistry1.inp) which contains the expressions (standard and extra) and the actual values of the model parameters. PEST uses three input files: (i) a template file (*.tpl e.g. chemistry1.tpl), identical to the ORCHESTRA chemistry file in which the parameters to be optimized are written in between pound signs # within the expressions and model equations. Within each iteration step, PEST writes the chemistry file in which the variable between the pound signs are replaced with the updated estimates of the parameter values; (ii) an instruction file (*.ins) that instructs PEST where and how to read the ORCHESTRA output from each model simulation; (iii) a control file (*.pst)

which brings it all together and gives the instructions for PEST to perform the parameter fitting, i.e., names of the template and instruction files, names of the parameters to be optimized with their minimum and maximum values, observation data to which the model has to be fitted, options and criteria for the parameter optimization and instruction lines to run the model (in this case ORCHESTRA).

The results of the fitting procedure are written in two PEST output text files: (i) a .rec file that lists all steps of the optimization procedure and give the optimum parameters, together with a 95% confidence interval and (ii) a .res file that summarizes final results and gives for each point the experimental value, model final estimation and corresponding residuals.

Examples and details on these files are given in sections 3.3 to 3.5. More examples can be found in the PEST manual (found on the webpage <http://www.pesthomepage.org/Downloads.php>).

3 Implementation of the unconstrained procedure in PEST-ORCHESTRA

3.1 *ORCHESTRA chemistry file (chemitry1.inp)*

We adjusted the implementation of the NICA-Donnan model in ORCHESTRA by substituting the model constants with variable names in the equations within the chemistry file. Thereby, the actual values of these variables can be changed in the model fitting procedure. Model variables are defined at the beginning of the file together with their value (which can be changed) according to the ORCHESTRA syntax: e.g. for the maximum proton site density (Qmax1) and the affinity distribution parameter (p1) of the type 1 sites:

```
@Var: Qmax1 5.88 //
```

```
@Var: p1 1.0 //
```

We then replaced the default values by these variable names within the equations:

```
@nicasite(HA1, HA, HA_don, p1, Qmax1 )
```

For the complete set of added definitions see below; the specific additions/modifications made to the chemistry file for the parameter optimization of proton binding to HA are marked in grey.

For proton titration modelling, p_1 and p_2 were fixed at 1.0, and $n_{H,1}$ and $n_{H,2}$, corresponding to the apparent heterogeneities m_1 and m_2 , were fitted.

To optimize the electrostatic parameter for the Donnan volume Vd calculation, we added the following extra-expression:

```
@Var: Vd 1.5 // declaration of the variable in ORCHESTRA
```

```
@Stage: (1, "Vd=10^(-b*log10(I)+b-1)") // calculation of Vd
```

Stage 1 is the first stage in the computational scheme. For an explanation of the computational stages in ORCHESTRA, see the original paper by Meeussen^[1].

The observational data to which the NICA-Donnan model is to be fitted by optimization of the model parameters is usually expressed in terms of humic charge (in mol·kg⁻¹), expressed in ORCHESTRA by the variable $H+.HA_{part}$.

The number 5 relates to the stage of the computations in which the expression is evaluated.

```
//***** Version: 9 Februari 2016 13:49 *****
@logactivities: // Indicates that this version uses log activities for entities.
(name.logact)

// Adding @forceLogKreactions: here will force the graphical editor to rewrite all
reactions in log K format.
// Adding @forceLinKreactions: here will force the graphical editor to rewrite all
reactions in linear K format.
// The graphical editor will not change the original format of reactions in inputfile
otherwise.
@NrDigits: 12 // number of decimal places in reaction coefficients.

//***** The database file(s) *****
// This section list the database files that are used by the interactive chemistry editor
// Syntax for local files: @database: minteq.txt
// Syntax for files on the internet: @database: www.meeussen.nl/orchestra/minteqv4.txt

@database: ../../bin/minteqv4.txt
@database: ../../bin/adsmodels.txt
//***** End of the database file(s) *****

//***** Generate output variables *****
//***** End of output variables *****

//***** The variables *****
@globalvar: H2O.logact 0
@Var: DHA_kg1 5.0E-4 //
@Var: H+.logact 1.0 // **
@Var: Qmax1 3.24 // proton site density of type-1 sites
@Var: Qmax2 2.88 // proton site density of type-2 sites
@Var: Vd 1.5 // Donnan volume (L/kg)
@Var: b 0.57 // b par. in relation Vd with ionic strength(Eq 5)
@Var: logKH1 2.66 // log proton binding constant of type-1 sites
@Var: logKH2 6.9 // log proton binding constant of type-2 sites
```



```

@Var: nH1 0.81 // non-ideality of type-1 sites
@Var: nH2 0.29 // non-ideality of type-2 sites
@Var: p1 1.0 // affinity distribution of type-1 sites
@Var: p2 1.0 // affinity distribution of type-2 sites
@Var: pH NaN //
@Var: watervolume 1.0 //
//***** End of the variables *****

//***** The extra expressions *****
@Stage: (1, "Vd=10^(-b*log10(I)+b-1)") // calculation of Vd
//***** End of the extra expressions *****

//***** The phases *****
@phase(tot)
@phase(liter)
@link_phase(liter, tot, "watervolume")
@phase(diss)
@link_phase(diss, liter, "1")
@phase(DHA_part_1)
@link_phase(DHA_part_1, diss, "1")
@phase(HA_part)
@link_phase(HA_part, DHA_part_1, "DHA_kg1")
//***** End of the phases *****

@include: uiobjects.txt
@globalvar: I .1
@davies()

//***** The primary entities *****
@Global: pH
@Var: H+.logact 1 **
@Calc:(1, "H+.logact = -pH")
@species(H+, 1)
@primary_entity(H+, pH, -9.0)
@entity(H2O, diss, 55.6 )
@primary_entity(H2O, 0.0)
@species(K+, 1)
@primary_entity(K+, -9.0, diss, 0.01)
@species(NO3-, -1)
@primary_entity(NO3-, -9.0, diss, 0.01)
//*****

//***** The entities *****
@species(OH-, -1)
@reaction(OH-, 1.00693e-14 , -1.0, H+, 1.0, H2O)

//***** The minerals *****
// This version can use external mineral files.
// Syntax: @external_mineral_file: name.txt
//***** End of the minerals *****

//***** Adsorption Models *****

//***** Adsorption Model HA *****

@nicamodel(HA, HA_part, 1, "Vd")

/** Automatically generated donnan species for HA
@donnanspecies(H+_HA_don, HA_don, H+, 1)
@donnanspecies(K+_HA_don, HA_don, K+, 1)
@donnanspecies(NO3-_HA_don, HA_don, NO3-, -1)
@donnanspecies(OH-_HA_don, HA_don, OH-, -1)

@nicasite(HA1, HA, HA_don, p1, Qmax1)
@nicaspecies(HA1-H, HA1, H+_HA_don, nH1, nH1, logKH1 )

@nicasite(HA2, HA, HA_don, p2, Qmax2)
@nicaspecies(HA2-H, HA2, H+_HA_don, nH2, nH2, logKH2 )
//***** End of Adsorption Model HA *****

//***** End of Adsorption Models *****

```

3.2 ORCHESTRA input file (input.dat)

Var: pH K+.diss NO3-.diss DHA_kg1

```
Data: 3.661564626 0.0091      0.0091      0.000123873
Data: 3.668367347 0.009102665 0.009102665 0.000123842
Data: 3.692176871 0.009110652 0.009110652 0.00012375
Data: 3.769557823 0.009134543 0.009134543 0.000123475
Data: 3.851190476 0.009155955 0.009155955 0.000123229
Data: 3.931122449 0.009174127 0.009174127 0.000123019
Data: 4.008503401 0.009189878 0.009189878 0.000122838
[...]
```

For the complete input.dat, see the website.

3.3 *PEST chemistry file (chemistry1LFA.tpl)*

The template file chemistry1LFA.tpl is used by PEST to generate the chemistry1.inp file for ORCHESTRA. The variables declared in PEST (see the *.pst file) are between hashtags and are to be replaced by their actual values in each iteration by PEST (e.g. #b #). For detailed information, see the PEST manual (fifth edition) section 3.2.

```
ptf #
//***** Version: 9 Februari 2016 13:49 *****
@logactivities: // Indicates that this version uses log activities for entities.
(name.logact)

// Adding @forceLogKreactions: here will force the graphical editor to rewrite all
reactions in log K format.
// Adding @forceLinKreactions: here will force the graphical editor to rewrite all
reactions in linear K format.
// The graphical editor will not change the original format of reactions in inputfile
otherwise.
@NrDigits: 12 // number of decimal places in reaction coefficients.

//***** The database file(s) *****
// This section list the database files that are used by the interactive chemistry editor
// Syntax for local files: @database: minteq.txt
// Syntax for files on the internet: @database: www.meeussen.nl/orchestra/minteqv4.txt

@database: ../../bin/minteqv4.txt
@database: ../../bin/adsmodels.txt
//***** End of the database file(s) *****

//***** Generate output variables *****
//***** End of output variables *****

//***** The variables *****
@globalvar: H2O.logact 0
@Var: DHA_kgl 5.0E-4 //
@Var: H+.logact 1.0 // **
@Var: Qmax1 3.24 //
@Var: Qmax2 2.88 //
@Var: Vd 1.5 //
@Var: b 0.57 //
@Var: logKH1 2.66 //
@Var: logKH2 6.9 //
@Var: nH1 0.81 //
@Var: nH2 0.29 //
@Var: p1 1.0 //
@Var: p2 1.0 //
@Var: pH NaN //
@Var: watervolume 1.0 //
//***** End of the variables *****

//***** The extra expressions *****
@Stage: (1, "Vd=10^(-#b #*log10(I)+#b #-1)")
//***** End of the extra expressions *****
```

```
//***** The phases *****
@phase(tot)
@phase(liter)
@link_phase(liter, tot, "watervolume")
@phase(diss)
@link_phase(diss, liter, "1")
@phase(DHA_part_1)
@link_phase(DHA_part_1, diss, "1")
@phase(HA_part)
@link_phase(HA_part, DHA_part_1, "DHA_kgl")
//***** End of the phases *****

@include: uiobjects.txt
@globalvar: I .1
@davies()

//***** The primary entities *****
@Global: pH
@Var: H+.logact 1 **
@Calc:(1, "H+.logact = -pH")
@species(H+, 1)
@primary_entity(H+, pH, -9.0)
@entity(H2O, diss, 55.6 )
@primary_entity(H2O, 0.0)
@species(K+, 1)
@primary_entity(K+, -9.0, diss, 0.01)
@species(NO3-, -1)
@primary_entity(NO3-, -9.0, diss, 0.01)
//*****

//***** The entities *****
@species(OH-, -1)
@reaction(OH-, 1.00693e-14 , -1.0, H+, 1.0, H2O)

//***** The minerals *****
// This version can use external mineral files.
// Syntax: @external_mineral_file: name.txt
//***** End of the minerals *****

//***** Adsorption Models *****

//***** Adsorption Model HA *****
@nicamodel(HA, HA_part, 1, "Vd")

/** Automatically generated donnan species for HA
@donnanspecies(H+_HA_don, HA_don, H+, 1)
@donnanspecies(K+_HA_don, HA_don, K+, 1)
@donnanspecies(NO3-_HA_don, HA_don, NO3-, -1)
@donnanspecies(OH-_HA_don, HA_don, OH-, -1)

@nicasite(HA1, HA, HA_don, p1, #Qmax1 # )
@nicaspecies(HA1-H, HA1, H+_HA_don, #nH1 #, #nH1 #, #logKH1 # )

@nicasite(HA2, HA, HA_don, p2, #Qmax2 # )
@nicaspecies(HA2-H, HA2, H+_HA_don, #nH2 #, #nH2 #, #logKH2 # )
//***** End of Adsorption Model HA *****
//***** End of Adsorption Models *****
```

3.4 PEST control file (H_LFA.pst)

The PEST control file gives the instructions for PEST to do the parameter fitting: names of the template and instruction files, parameters to be optimized and their ranges, values of the observations and instruction lines to run the model (in this case ORCHESTRA). For explanation see PEST manual section 4.

pcf

```
* control data
norestart estimation
7 255 4 0 1 //number of: parameters, observation data, parameters
groups, prior information, observation groups
1 1 single point 1 0 0
5.0 2.0 0.3 0.01 7
2.0 2.0 0.001
0.1
30 0.005 4 4 0.01 4
1 1 1
* parameter groups
don relative 0.01 0.0 switch 1.5 outside_pts
logk relative 0.01 0.0 switch 1.5 outside_pts
n relative 0.01 0.0 switch 1.5 outside_pts
charge relative 0.01 0.0 switch 1.5 outside_pts
* parameter data
b none relative 0.28 0.1 0.8 don 1.0 0.0 1
logKH1 none relative 3.21 1 4.5 logk 1.0 0.0 1
logKH2 none relative 8.81 6 9.5 logk 1.0 0.0 1
nH1 none relative 0.55 0.1 1.0 n 1.0 0.0 1
nH2 none relative 0.20 0.1 1.0 n 1.0 0.0 1
Qmax1 none relative 3.21 0.1 9 charge 1.0 0.0 1
Qmax2 none relative 6.64 0.1 9 charge 1.0 0.0 1
* observation groups
group_1
* observation data
Q1 -1.217541702 1 group_1
Q2 -1.227482645 1 group_1
Q3 -1.245744916 1 group_1
[...]
Q253 -7.120110204 1 group_1
Q254 -7.159708336 1 group_1
Q255 -7.188094845 1 group_1
* model command line
orchestra_pest.bat
* model input/output
chemistry1LFA.tpl chemistry1.inp
outputFLFA.ins output_pest.dat
```

3.5 *PEST instruction file (outputFLFA.ins)*

The instruction file instructs PEST how to read the model output file (ORCHESTRA output file). For more details, see PEST manual section 3.3.

```
pif *
*Var:*
l1 *Data:* w w w w w w w w w w
w w w w !Q1!
l1 *Data:* w w w w w w w w w w
w w w w !Q2!
l1 *Data:* w w w w w w w w w w
w w w w !Q3!
[...]
```

11	*Data:*	w	w	w	w	w	w	w	w	w	w
	w	w	w	w	!Q253!						
11	*Data:*	w	w	w	w	w	w	w	w	w	w
	w	w	w	w	!Q254!						
11	*Data:*	w	w	w	w	w	w	w	w	w	w
	w	w	w	w	!Q255!						

4 Implementation of the master curve procedure in PEST-ORCHESTRA

4.1 Description of the master curve approach

With the master curve procedure only the electrostatic part of the model, the Donnan model is used. The Donnan volume, which is assumed constant at a certain ionic strength, is optimized for each ionic strength such that the proton titration curves at different ionic strength merge into a single master curve. This can be done by relating the humic charge of the HA (Q) to the negative logarithm of the proton concentration in the Donnan volume (pH_{don}). With the Donnan model the concentrations of protons and other cat- and anions in the Donnan volume are calculated such that the negative charge of the humic acid is counterbalanced by the surplus of positive charge in the Donnan volume. These calculations are done for similar series of humic charge at each ionic strength. For each point the input for the calculation consists of: the humic charge, the pH associated to that charge and the concentrations of the cat- and anions of the background electrolyte. In order to have the input for the three different ionic strengths for similar values of humic charge we derived relations between Q and pH for each ionic strength. These relations were then used to calculate the pH associated to a certain Q at each ionic strength. For this purpose we used a 6th order polynomial relation with pH as a function of Q . For each ionic strength we constructed an input file with 100 points of humic charge (equidistant). To stay within the boundaries of the experimental values the lowest charge to be used is the maximum of the lowest charge of the 3 titration curves and the highest charge to be used is the minimum of the highest charges of the 3 titration curves. The construction of the Q -pH input data can be found in the Excel worksheet FH-23-MC.xls on the website.

The template for the chemistry file (chemistry1.tpl) with the implementation of the Donnan model is listed below (Paragraph 4.2). For each ionic strength (in this case 3 datasets) the proton concentrations in the Donnan volume are calculated for a similar series of humic charge. For each ionic charge there is a separate chemistry-, input- and output file (*i.e.*, “chemistryX.inp”, “inputX.dat” (Paragraph 4.3) and “outputX.dat”, with X=1,2,3). The proton concentrations in the Donnan volume are calculated for each point of humic charge for the three ionic strengths and their values (pHdonX with X=1,2,3 for the 3 different ionic strengths) are written to the three output files (“outputX.dat”). Thereafter ORCHESTRA reads the values of pHdonX from the 3 output files and calculates dif1 and dif2, being equal to pHdon1 - pHdon2 + 1 and pHdon1 - pHdon3 + 1, respectively. The instructions for this calculation are in the file “postprocess.inp” (Paragraph 4.4). The values of dif1 and dif2 are written to the file outputpost.dat which is the outputfile to be read by PEST according to the format of the outputfile described in “outputP.ins”. PEST optimizes the model parameters by minimizing the difference of dif1 and dif2 with the value of the observations (which is set to 1 for all cases). The value of 1 (and the addition of 1 in the calculation of dif1 and dif2) was chosen instead a value of 0 to enable the calculation of r^2 by PEST.

The instructions for ORCHESTRA are defined in the file composer.xml (see the webpage).

The titration curves at the various ionic strengths are fitted to a polynomial of degree 6 relating the pH to $\log Q$. Thereby, we are able to build titration curves with the same Q values for the different ionic strengths. The proton concentration in the Donnan volume $H+don$ (defined as a variable of the system at the beginning of the chemistry file) is calculated at all pH and I values in ORCHESTRA using equation (3) of the manuscript **Erreur! Source du renvoi introuvable**.applied to protons:

@Stage: (5, "-log(H+don) = -log10({H+.con}) -{HA_don.logact}"),

Where $\{H+.con\}$ is the proton concentration in the bulk solution and $\{HA_{don}.logact\}$ is the Boltzman factor $\exp(-\frac{z_H F \Psi_D}{RT})$.

4.2 *PEST chemistry file (chemistry1.tpl)*

```
ptf #
//***** Version 25 May 2011 10:19 *****
@logactivities: // Indicates that this version uses log activities for entities.
(name.logact)

// Adding @forceLogKreactions: here will force the graphical editor to rewrite all
reactions in log K format.
// Adding @forceLinKreactions: here will force the graphical editor to rewrite all
reactions in linear K format.

//***** The database file(s) *****
// This section list the database files that are used by the interactive chemistry editor
// Syntax for local files: @database: minteq.txt
// Syntax for files on the internet: @database: www.meeussen.nl/orchestra/minteqv4.txt

@database: ../../bin/minteqv4.txt
@database: ../../bin/adsmodels.txt
//***** End of the database file(s) *****

//***** Generate output variables *****
//***** End of output variables *****

//***** The variables *****
@globalvar: H2O.logact 0
@Var: DHS_kgl 1.0E-6 //
@Var: H+.logact 1.0 // **
@Var: Q -0.9 //
@Var: pH NaN //
@Var: pHdon1 4.0 //
@Var: adon #adon # // alpha parameter in equation donnan volume
@Var: bdon #bdon # // beta parameter in equation donnan volume
@Var: Vd 3 // donnan volume
@Var: watervolume 1.0 //
//***** End of the variables *****

//***** The extra expressions *****
//* The format of extra expressions is:
//* Stage:(1,"expression")
//* (Stage: is used instead of Calc: here to prevent interpretation by the GUI.)
//* For the calculations Stage: and Calc: are equivalent.
@Stage: (1, "Vd=10^(adon+bdon*log10(I))")
@Stage: (5, "pHdon1=-log10({H+.con})-{{HS_don.logact}}") // {{HS_don.logact}} being the Boltzman
factor
//***** End of the extra expressions *****

//***** The phases *****
@phase(gas)
@phase(tot)
@phase(liter)
@link_phase(liter, tot, "watervolume")
@phase(diss)
@link_phase(diss, liter, "1")
@phase(DHS_part_1)
@link_phase(DHS_part_1, diss, "1")
@phase(HS_part)
@link_phase(HS_part, DHS_part_1, "DHS_kgl")
//***** End of the phases *****

// The uiobjects1 class can be used to add user defined content.
// It is literally read and written by the GUI.
```

```

@class: uiobjects1(){%
  @include: uiobjects.txt
}%
@uiobjects1()

@globalvar: I .1
@davies()

//***** The primary entities *****
@species(Cl-, -1)
@primary_entity(Cl-, -9.0, diss, 1.0E-9)
@Global: pH
@Var: H+.logact 1 **
@Calc:(1, "H+.logact = -pH")
@species(H+, 1)
@primary_entity(H+, pH, -9.0)
@entity(H2O, diss, 55.6 )
@primary_entity(H2O, 0.0)
@species(K+, 1)
@primary_entity(K+, -9.0, diss, 1.0E-9)
//*****

//***** The entities *****

@species(OH-, -1)
@reaction(OH-, 1.00693e-14 , -1.0, H+, 1.0, H2O)

@class: extra_entities(){%
  // Here you can put some arbitrary extra code.
}%
@extra_entities()

//***** The minerals *****
// This version can use external mineral files.

// Syntax: @external_mineral_file: name.txt
//***** End of the minerals *****

//***** Adsorption Models *****

//***** Adsorption Model HS *****

// Note that the surface charge of this model is in eq/kg
@nicamodel(HS, HS_part, 1, Vd,Q)

/** Automatically generated donnan species for HS
@donnanspecies(Cl-_HS_don, HS_don, Cl-, -1)
@donnanspecies(H+_HS_don, HS_don, H+, 1)
@donnanspecies(K+_HS_don, HS_don, K+, 1)
@donnanspecies(OH-_HS_don, HS_don, OH-, -1)
//***** End of Adsorption Model HS *****

//***** End of Adsorption Models *****
//***** add @userfriendlyformat: to get ouput of reactions in user friendly format *****

```

4.3 *Input file (input1.dat)*

Header and first three records are given here, the complete file can be found on the website.

Var:	pH	Cl-.diss	K+.diss	DHS_kg1	Q
Data:	4.1587	0.009	0.009	1.25E-04	-1.63E+00
Data:	4.2107	0.009	0.009	1.25E-04	-1.68E+00
Data:	4.2610	0.009	0.009	1.25E-04	-1.73E+00

4.4 *postprocess.inp*

```
@Var: pHdon1 3.0
@Var: pHdon2 3.0
@Var: pHdon3 3.0
@Var: dif1 10
@Var: dif2 10
@Calc: (1, "dif1 = 1 + pHdon1 - pHdon2" )
@Calc: (1, "dif2 = 1 + pHdon1 - pHdon3" )
```

5 Exclusion of electrostatic interactions in the NICA-Donnan model

The content below was added to the ORCHESTRA chemistry file “chemistry1.inp” to overwrite the standard object for the Donnan model (linked to the NICA model) as it is defined in the object file “objects2016.txt”. The content was added immediately after the phase definitions.

```
// The uiobjects1 class can be used to add user defined content.
// It is literally read and written by the GUI.
@class: uiobjects1(){%
    @include: ..\..\bin\objects2016.txt
    @Class: nicamodel(name, parentphase, concentration, donnanVolume){
    // define the nica surface phase
    @phase(<name>, <parentphase>, <concentration>)

    //-----
    // Add a donnan phase to this surface, which consist of a phase,
    linked to the
    // surface via volume/kg. The donnan phase is also an entity (act =
    boltzman factor, sum = charge balance)
    // unknown is activity, equation is charge balance at surface = 0
    //-----
    @Var: <name>donvol 1.5 // The Donnan volume + default value
    //@Calc:(1,"<name>donvol = <donnaVolume>")
    @Calc:(1,"<name>donvol = 0")
    @phase(<name>_don, <name>, <name>donvol)
    @GlobalVar: <name>_don.logact 0
    @entity(<name>_don, <name>_don, 0)
    //@Uneq2: unknown:(name:, <name>_don.logact, delta:, 1e-6, type:,
    lin, step: , 1, default:, 0, iia:, true) equation:(name:,
    <name>_don.<name>, tol: , 1e-4)
    //-----
    }
    %}
@uiobjects1()
```

6 Implementation of metal-binding parameters optimization

Here, examples of modifications of the ORCHESTRA chemistry file in the case of metal binding to HS are given using Cd as the metal of interest. The same procedure was applied with Zn. The new variables ($n_{i,1}$, $n_{i,2}$, $\log K; \tilde{~}_{i,1}$, $\log K; \tilde{~}_{i,2}$) were defined at the beginning of the chemistry file, and the corresponding NICA-Donnan equations were modified accordingly:

```
@nicaspecies(HA1-Cd, HA1, Cd+2_HA_don, nCd1, nH1, logkCd1)
```

```
@nicaspecies(HA2-Cd, HA2, Cd+2_HA_don, nCd2, nH2, logkCd2)
```

Values of $n_{H,1}$ and $n_{H,2}$ can be calculated from m_1 and m_2 determined from proton titrations fitting according to the following expressions:

```
@Stage: (1, "nH1=m1/p1")
```

```
@Stage: (1, "nH2=m2/p2")
```

Input variables are the pH, electrolyte concentrations ($Na^+.diss$, $NO_3^-.diss$), LFA concentration (in $kg \cdot L^{-1}$) and free Cd^{2+} and Zn^{2+} activities ($Cd+2.logact$, $Zn2+.logact$). The parameters to be fitted during the optimization procedure were the amount of metal bound to the HS $\log CdHA$ and $\log ZnHA$ (in $mol \cdot kg^{-1}$). The AGNES method determines the free metal concentration. In that case, the bound fraction of metal corresponds to the fraction bound to the HS together with the inorganic metallic complexes, and it can be calculated as follows:

```
@Stage: (5, "logCdHA = log10(((Cd+2.DHA_part_1)+(CdNO3+.diss))/(DHA_kgl))")
```

```
@Stage: (5, "logZnHA = log10(((Zn+2.DHA_part_1)+(ZnNO3+.diss))/(DHA_kgl))")
```

The three dataset (H^+ -, Cd- and Zn-binding) were fitted simultaneously, and the input file had columns for the free Cd^{2+} and Zn^{2+} activities. Since ORCHESTRA does not use the activity directly, but instead the logarithm of the activity ($*.logact$), the software did not support a 0 input value. For the cases when one or the other metal ion was not present in the system, the metal activities were thus set at a default value of -50.

7 Results

Table S1. Summary of experimental data used in the parameter optimization

	I (M)	electrolyte	pH	LFA con. (mg/L)	Metal conc. (μ M)	Reference
H	0.01, 0.03, 0.1	KNO ₃	3.6-10.4	125	-	Pinheiro et al. ^[5]
Cd	0.01	NaNO ₃	6, 7, 8	9 - 32	0.1 - 5	This study
Zn	0.01	NaNO ₃	6, 7, 8	30 - 65	0.06 - 6.10	This study

7.1 Results of proton-binding parameters optimization

7.1.1 Unconstrained procedure

Table S2. Parameter correlation coefficient matrix obtained for the unconstrained optimization using parameter b

	b	$\log \tilde{K}_{H,1}$	$\log \tilde{K}_{H,2}$	$n_{H,1}$	$n_{H,2}$	$Q_{\max,1}$	$Q_{\max,2}$
b	1.000	0.2338	0.1485	-0.1212	0.1237	0.1340	-0.1054
$\log \tilde{K}_{H,1}$	-	1.000	0.2409	0.8023	-0.9172	-0.8329	0.9183
$\log \tilde{K}_{H,2}$	-	-	1.000	-0.2381	-0.0992	0.2012	0.2242
$n_{H,1}$	-	-	-	1.000	-0.9257	-0.9899	0.8826
$n_{H,2}$	-	-	-	-	1.000	0.9525	-0.9896
$Q_{\max,1}$	-	-	-	-	-	1.000	-0.9087
$Q_{\max,2}$	-	-	-	-	-	-	1.000

Table S3. Parameter correlation coefficient matrix obtained for the unconstrained optimization using parameters α and β

	α	β	$\log \tilde{K}_{H,1}$	$\log \tilde{K}_{H,2}$	$n_{H,1}$	$n_{H,2}$	$Q_{\max,1}$	$Q_{\max,2}$
α	1.000	0.9041	0.4161	0.1264	-0.1486	0.0992	0.0993	-0.1150
β	-	1.000	0.3642	0.0747	-0.0826	0.0436	0.0390	-0.0622
$\log \tilde{K}_{H,1}$	-	-	1.000	0.4174	0.7208	-0.8302	-0.7736	0.8263
$\log \tilde{K}_{H,2}$	-	-	-	1.000	-0.0279	-0.2216	-0.0197	0.2933
$n_{H,1}$	-	-	-	-	1.000	-0.9593	-0.9923	0.9420
$n_{H,2}$	-	-	-	-	-	1.000	0.9780	-0.9957
e	-	-	-	-	-	-	1.000	-0.9599
$Q_{\max,2}$	-	-	-	-	-	-	-	1.000

7.1.2 Master curve procedure

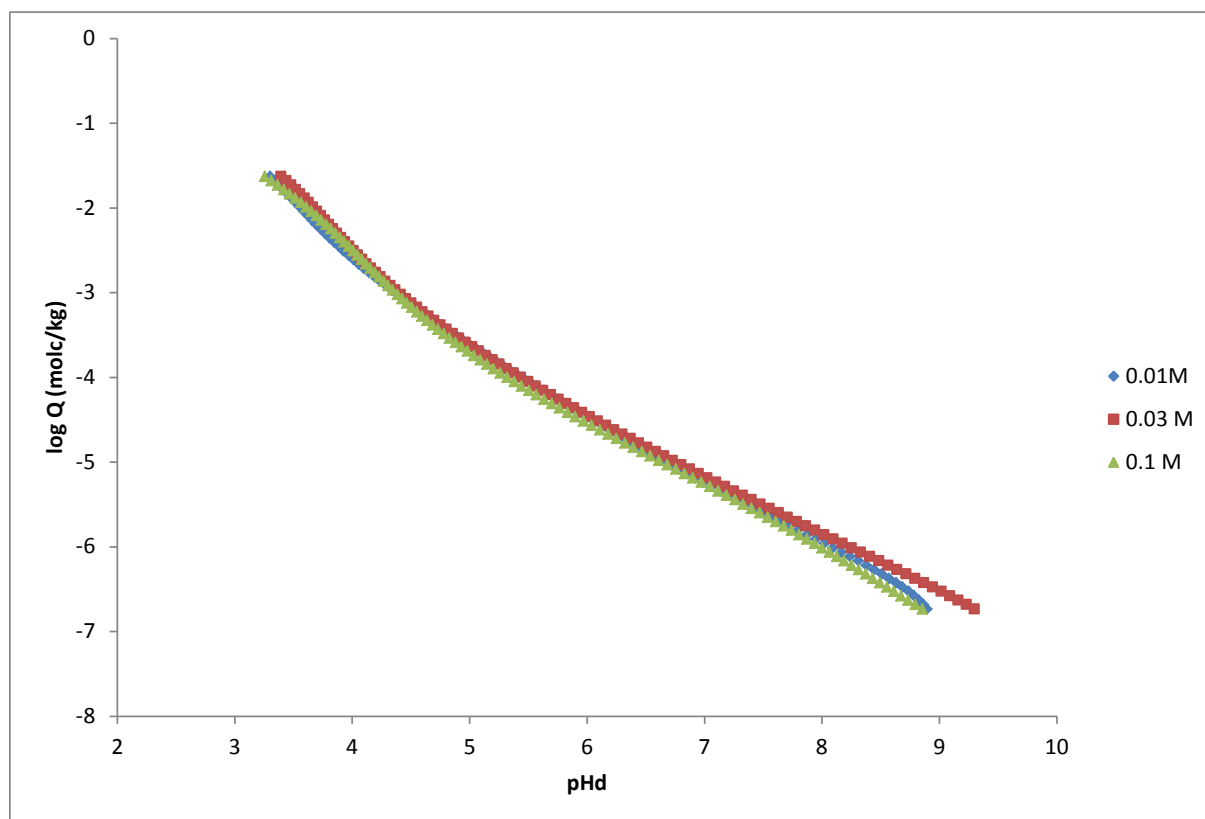


Figure S3. “Master Curve” with logQ as a function of pHd ($=-\log[H^+]$ donnan).

Table S4. Parameter correlation coefficient matrix obtained for the first step of the master curve procedure: Optimization of α and β parameters

	α	β
α	1.000	0.8753
β	-	1.000

Table S5. Parameter correlation coefficient matrix obtained for the second step of the master curve procedure: Optimization of other parameters

	$Q_{\max,1}$	$Q_{\max,2}$	$\log \tilde{K}_{H,1}$	$\log \tilde{K}_{H,2}$	$m_{H,1}$	$m_{H,2}$
$Q_{\max,1}$	1.000	-0.9885	0.4444	-0.7533	-0.9903	0.9923
$Q_{\max,2}$	-	1.000	-0.3228	0.8440	0.9676	-0.9990
$\log \tilde{K}_{H,1}$	-	-	1.000	0.1748	-0.5168	0.3376
$\log \tilde{K}_{H,2}$	-	-	-	1.000	0.6969	-0.8269
$m_{H,1}$	-	-	-	-	1.000	-0.9713
$m_{H,2}$	-	-	-	-	-	1.000

7.1.3 2-step procedure

Table S6. Parameter correlation coefficient matrix obtained for the first step of the 2-step procedure: Optimization of $Q_{\max,i}$

	$n_{H,1}$	$n_{H,2}$	$Q_{\max,1}$	$Q_{\max,2}$
$n_{H,1}$	1.000	0.2782	-0.3487	0.0749
$n_{H,2}$	-	1.000	0.6817	-0.7608
$Q_{\max,1}$	-	-	1.00	-0.8936
$Q_{\max,2}$	-	-	-	1.000

Table S7. Parameter correlation coefficient matrix obtained for the second step of the 2-step procedure: Optimization of other parameters

	$\log \tilde{K}_{H,1}$	$\log \tilde{K}_{H,2}$	$n_{H,1}$	$n_{H,2}$	alpha	beta
$\log \tilde{K}_{H,1}$	1.000	0.9689	-0.5352	-0.2924	0.9635	0.7854
$\log \tilde{K}_{H,2}$	-	1.000	-0.4774	-0.2693	0.9565	0.7886
$n_{H,1}$	-	-	1.000	0.6467	-0.5453	-0.4808
$n_{H,2}$	-	-	-	1.000	-0.2668	-0.2657
alpha	-	-	-	-	1.000	0.9123
beta	-	-	-	-	-	1.000

7.2 Comparison with generic Cu-binding parameters derived by Milne et al. (2003)

In order to test whether the optimization using the combination of PEST-ORCHESTRA gives similar results for the NICA-Donnan parameters as the optimization by Milne et al. (2003)^[4] using FIT^[6] we optimized the generic parameters for Cu binding to FA. The data consists of 11 data sets containing 541 adsorption data, exactly similar as to those used by Milne et al. (2003)^[4]. The NICA-Donnan parameters $\log \tilde{K}_{Cu,1}$, $\log \tilde{K}_{Cu,2}$, $n_{Cu,1}$ and $n_{Cu,2}$ were optimized simultaneously without any constraints. The NICA-Donnan parameters b , p_1 and p_2 were set to their generic values^[4]. Input data for ORCHESTRA were pH, the free ion activity of Cu^{2+} and the solution concentration of the background electrolyte. Parameters were optimized to get an optimal fit between the Cu bound to FA calculated by the NICA-Donnan model and the experimentally derived data. The results using PEST-ORCHESTRA are nearly similar to those using the FIT algorithm (see Table S8).

Table S8. Comparison between generic NICA-Donnan parameters optimised using the PEST-ORCHESTRA modelling software with those optimised using the FIT software

	PEST-ORCHESTRA	FIT ^[4]	95% confidence PEST-ORCHESTRA
$\log \tilde{K}_{\text{Cu},1}$	0.27	0.26	0.18 - 0.35
$\log \tilde{K}_{\text{Cu},2}$	8.17	8.26	7.84 - 8.51
$n_{\text{Cu},1}$	0.56	0.53	0.54 - 0.58
$n_{\text{Cu},2}$	0.36	0.36	0.33 - 0.39
r^2	0.96	0.90	
RMSE	0.17	0.19	

References

- [1] J. C. L. Meeussen, ORCHESTRA: An object-oriented framework for implementing chemical equilibrium models. *Environ. Sci. Technol.* **2003**, *37*, 1175–1182.
- [2] T. Hiemstra, W. H. van Riemsdijk, A surface structural approach to ion adsorption: the charge distribution (CD) model. *J. Colloid Interface Sci.* **1996**, *179*, 488–508.
- [3] C. J. Milne, D. G. Kinniburgh, E. Tipping, Generic NICA-Donnan model parameters for proton binding by humic substances. *Environ. Sci. Technol.* **2001**, *35*, 2049–2059.
- [4] C. J. Milne, D. G. Kinniburgh, W. H. van Riemsdijk, E. Tipping, Generic NICA-Donnan model parameters for metal-ion binding by humic substances. *Environ. Sci. Technol.* **2003**, *37*, 958–971.
- [5] J. P. Pinheiro, A. M. Mota, M. F. Benedetti, Lead and calcium binding to fulvic acids: salt effect and competition. *Environ. Sci. Technol.* **1999**, *33*, 3398–3404.
- [6] D. G. Kinniburgh, *FIT User Guide, Technical Report WD/93/23* **1993** (Keyworth).