



**V. N. Karazin**  
**Kharkov**  
**National**  
**University**  
*(since 1804)*

# **Quantitative Physico-Chemical Analysis of Equilibria on Surfaces of Functionalized Materials: Problems, Achievements and Data Analysis**

**Yuriy Kholin**



# **Department of Technical Chemistry**

**1804-1834:**      **Department of Trade and  
Technology**

**1835-1863:**      **Department of Technology,  
Agriculture and Architecture**

**Since 1863:**      **Department of Technical  
Chemistry**

## Lecture Courses:

- Chemical Informatics and Chemometrics
- Simulation of Processes in Chemistry and Chemical Engineering
- Modern Achievements in Chemical Kinetics
- Chemical Ecology
- High-Molecular Compounds
- Organic Chemistry (for students of biologists)
- 7 courses for students specialized in design of materials and chemometrics



## Research Activity:

- Data analysis in chemistry and chemometrics
- Quantitative physico-chemical analysis
- Physical and coordination chemistry of functionalized materials
- Preconcentration and test analysis
- Mathematical and quantum chemistry
- Heterogeneous kinetics
- Organic luminophores

# Topics to be discussed:

1. What is the quantitative physico-chemical analysis (QPCA)? Who and why has a need to perform the QPCA study?
2. Main steps of the QPCA. The ill-posed nature of the computational step.
3. Two types of the energetic heterogeneity and distinguishing them. The numerical estimation of the biographic heterogeneity.
4. The data analysis methods applied to handle the primary experimental data of the QPCA.
5. Some results of exploring surfacely modified materials by the QPCA.

# 1. What is the quantitative physico-chemical analysis?

The quantitative physicochemical analysis is the method to determine the stoichiometric composition and stability constants of species in equilibrium systems by analyzing the composition–property dependencies.

**Composition-property dependency::**

$$A_k = \zeta(\lambda, n_{*k}), \quad (1)$$

A is the measured property (pH, absorbance, distribution coefficient, etc);

k is the number of the studied mixture (experimental point),  $1 \leq k \leq N$  ;

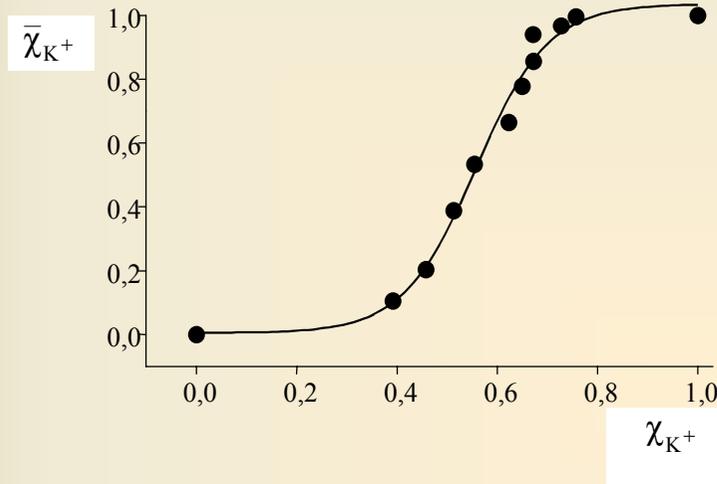
N is the number of points;

$\lambda$  is the analytical position (for example, a wavelength at the multiwave spectrophotometry;

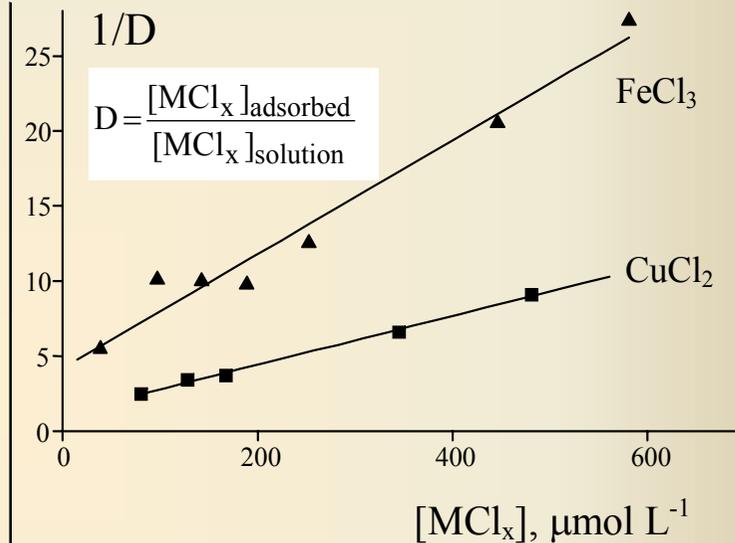
$n_{*k}$  are the initial quantities of reagents in the k-th mixture (known from the conditions of experiment);

$\zeta$  is a function.

## Examples of simple composition – property dependencies



The ion exchange isotherm of  $K^+$ .  
Plot of the  $K^+$  ion molar fractions in the ion exchanger phase ( $\bar{\chi}_{K^+}$ ) against its molar fractions in solution ( $\chi_{K^+}$ ).



Sorption of metal chlorides at the surface of chemically modified silica. Dependence of the inverse distribution coefficients against the equilibrium concentrations of the metal salts in solution.

*From the **single** composition – property dependency the following information may be obtained:*

- the number of species present in the systems under the equilibrium conditions;
- the stoichiometric compositions of species;
- the stability constants of species;
- the yields of each species dependent on the initial compositions of the systems;
- the intensity factors of the species (such as molar absorptivities).

Such information is valuable for choosing optimal conditions of separation, isolation, masking, determination of reagents.

The stability constants of transition metal complexes with amino-carbonic acids may be forecasted using regression models:

$$\lg \beta_{CuI} = -2,9 + 5,24n_{COO} + 1,9n_{OH} + 7,9n_N + 0,2n^2_{COO} - 0,6n_2 - 0,6n_N n_{COO} + \dots$$

Некоторые из коэффициентов уравнений, полученных из сведений о константах устойчивости для 20—30 °С и ионной силе 0,1 (приведены только коэффициенты перед значимыми факторами)

Фактор	Mn <sup>2+</sup>	Fe <sup>2+</sup>	Co <sup>2+</sup>	Ni <sup>2+</sup>	Cu <sup>2+</sup>	Zn <sup>2+</sup>	Pb <sup>2+</sup>	Cd <sup>2+</sup>
1	-1,6	-1,5	-1,4	-2,9	-2,9	-2,6	-1,7	-1,9
$n_{COO}$	2,7	3,1	2,6	3,9	4,6	3,9	4,0	2,4
$n_N$	2,7	3,2	4,1	5,9	7,9	4,8	3,5	4,2
$n_{OH}$	—	—	—	2,4	1,9	0,7	—	—
$n^2_{COO}$	—	—	—	—	-0,2	-0,2	-0,4	0,3
$n^2_N$	-0,2	-0,3	-0,3	-0,5	-0,6	-0,8	-0,3	-0,3
$n^2_{OH}$	—	—	—	-0,4	—	—	—	—
$n_{COO}n_N$	-0,3	-0,3	-0,4	-0,6	-0,6	-0,7	-0,4	-1,0
$n_{COO}n_{OH}$	—	—	—	—	-0,7	-0,3	—	—
$n_N n_{OH}$	—	—	-0,5	-0,8	-1,0	-0,7	-0,9	-0,4

The comparison of compositions and stabilities of species for related systems makes it possible also to do important conclusions about the non-thermodynamic properties of reagents. Hence, the QPCA supplements essentially the structure-sensitive structural methods.

## 2. The main steps of the QPCA

I. To choose the experimental method, the property(ies) to be measured and the initial concentrations of reagents. As a rule, the measured property may be presented as

$$A_{jk} = \sum_{i=1}^S \alpha_{ji} [L_i]_k,$$

where  $L_i$  are the reagents,  $[L_i]_k$  is the equilibrium concentration of  $L_i$  in the  $k$ th mixture,  $\alpha_{ji}$  is the intensity factor of  $L_i$  at the analytical position  $\lambda_j$ .

II. To execute the structural identification of the model: to find a such form of function  $\zeta$ , that allows to calculate the equilibrium concentrations  $[L_i]$  from the initial ones. In the simplest case, the structure of the model is assigned by the mass-action law and the conditions of the material balance. If reactions at the solid/liquid interface, ion-exchange processes and so on are considered, the structure of the model is to be complicated in an appropriate way.

**III.** To execute the parametric identification of the model: to find such values of parameters  $\theta$  that correspond to the "best" (in a certain sense) approximation of the experimental data:

$$|\theta\rangle = \arg \min U(\theta).$$

where  $U$  measures the goodness of fit.

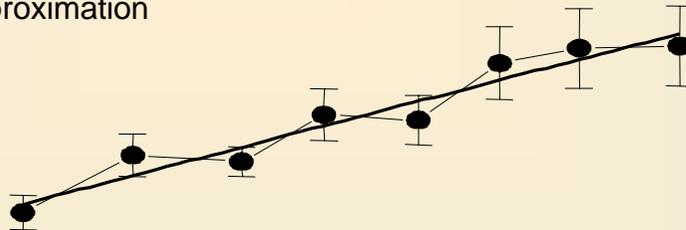
The main difficulties of the parametric identification arise from the fact that **the composition – property dependency is set by a table of values rather than an analytical expression**, whilst **the density of the experimental errors is unknown**.

The inevitable consequences are as follows:

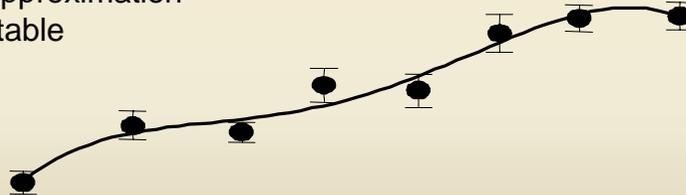
- a) there is no unique set of parameters  $\theta$  that fits the experimental data inside the limits of their errors.

From the chemical point of view, it means that redundant (non-existent) complexes with false stability constants may be detected;

if experimental errors are great  
the linear approximation  
is acceptable



if experimental errors are moderate  
the linear approximation  
is unacceptable



- a) If all measured values  $A_{lk}$  are independent and the density of their errors  $\varepsilon$  is normal (Gaussian), then maximum likelihood principle leads to the following criterion function for calculating unknown parameters  $\theta$ :

$$U(\theta) = \sum_{l=1}^L \sum_{k=1}^N w_{lk} \Delta_{lk}^2$$

where  $w_{lk}$  is the statistical weight corresponding to variance estimate  $\sigma^2(A_{lk})$ ,  $\Delta_{lk}$  is the discrepancy,

$$\Delta_{lk} = A_{lk}^{\text{calculated}} - A_{lk}^{\text{experimental}}$$

But we are never convinced that this supposition about normality is valid. On the contrary, at least 1/3 of experimental data does not correspond to the normal distribution!

***In this case, estimates obtained using the least squares method are biased, insignificant and inefficient.***

- b) The model adequacy criteria (such as  $\chi^2$ ) are more or less reliable also in the case of the normal distribution of experimental errors.

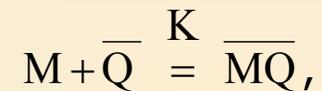
***The erroneous conclusion about the adequacy of a bad model may be easily made.***

## **Conclusions:**

- **the parametric identification belongs to the ill-posed problems, and so numerically stable (regularized) computational methods complemented with the comprehensive analysis of the possible redundancy of the model should be used;**
- **the usually used the least-squares estimates have to be verified by application of the robust methods insensible to the violation of the statistical assumptions;**
- **only applying several statistical adequacy criteria may ensure the reliable conclusions on the adequacy of the models;**
- **not standard (universal) but specially developed software programs should be used for calculations at the QPCA.**

### 3. Two types of the energetic heterogeneity and distinguishing them. The numerical estimation of the biographic heterogeneity

The formation of sorption complex  $\overline{MQ}$ :



where  $K$  is the affinity constant.

The **biographic heterogeneity**: the degree of occupation of binding sites by bonded species does not affect the scope of biographic heterogeneity.

The **evolutionary heterogeneity**: change of surface charge and electrostatic potential in dependence on the ion sorption degree.

To characterize the energetic heterogeneity it is necessary

- to measure the dependence of  $f([M])$  on  $[M]$ ;
- to postulate a model which allows to separate effects of biographic and evolutionary heterogeneities;
- to calculate numerically the model parameters. To perform this it is necessary to solve in respect to  $p(K)$  the linear first kind Fredholm equation:

$$f([M]) = \int_0^{\infty} \theta^{\text{local}}([M], K) \cdot p(K) dK,$$

where  $p(K)$  is the nonnegative differential distribution function of affinity constants  $K$ ;

the kernel of the integral equation  $\theta^{\text{local}}([M], K)$  is the local binding isotherm.

In the limiting case when both electrostatic effects and lateral interactions are negligible, the kernel  $\theta^{\text{local}}([M], K)$  reduces to the Langmuir isotherm

$$\theta^{\text{local}}([M], K) = \frac{K \cdot [M]}{1 + K \cdot [M]}.$$

This equation is the Fredholm integral equations of the first kind, and the search for its solution is the typical example of the ill-posed problem.

Our algorithm CAS deals with the integral distribution function and solves another integral equation

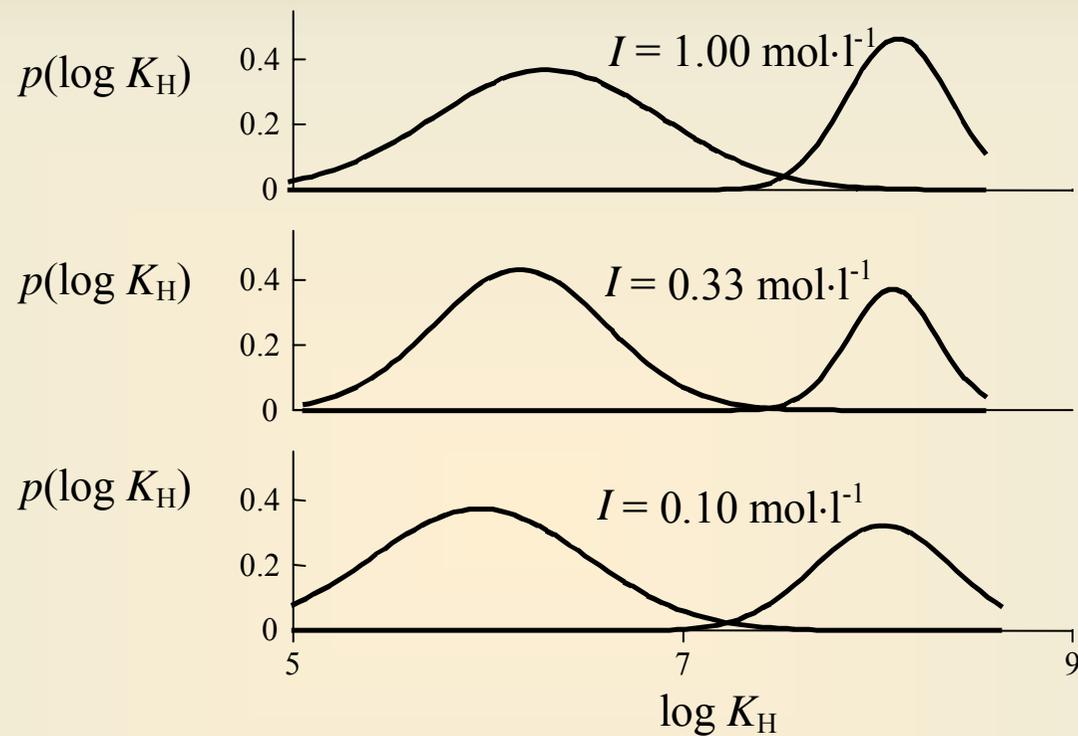
$$1 - \theta([M]) = \int_0^{\infty} P(K) \frac{[M]}{(1 + K \cdot [M])^2} dK$$

The main features of the method are:

- a) calculation of the integral distribution function  $P(K)$  on the compact set of restricted ( $0 \leq P(K) \leq 1$ ) and non-descending functions and
- b) using alpha-regularization technique to calculate the non-negative distribution function  $p(K)$ .

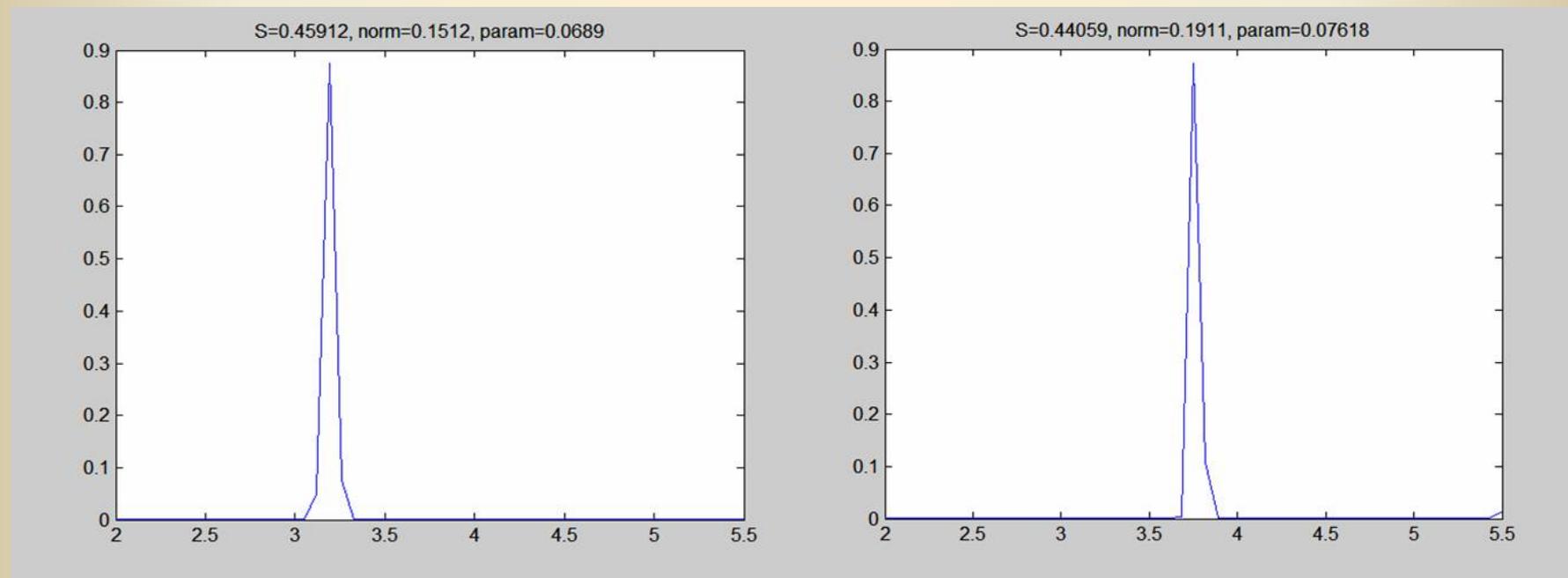
The full description is published in:

Kholin Yu., Myerniy S., Varshal G., *Adsorption Science & Technology*. 2000. **18**, 267-294.



The differential distributions of organic amines grafted on silica surface on logarithms of protonization constants

Estimation of the narrow distribution functions is the most complicated problem. The novel algorithm proposed by Dr S. Myerniy and based on the Shannon information theory solves this problem successfully.



# The data analysis methods applied to handle the primary experimental data of the QPCA

## The models based on the mass-action law

### The methods of the data analysis theory applied to avoid possible errors:

- cross-validation
- robust estimation of parameters
- elimination of the redundant species from the model on the base of the SVD decomposition of the Jacobi matrix, studying the correlation coefficients and the equilibrium concentrations (overcoming the ill-posed nature of the problem)

#### **Cross-validation:**

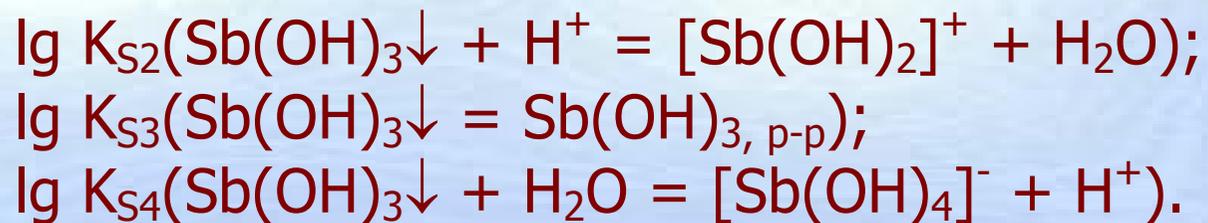
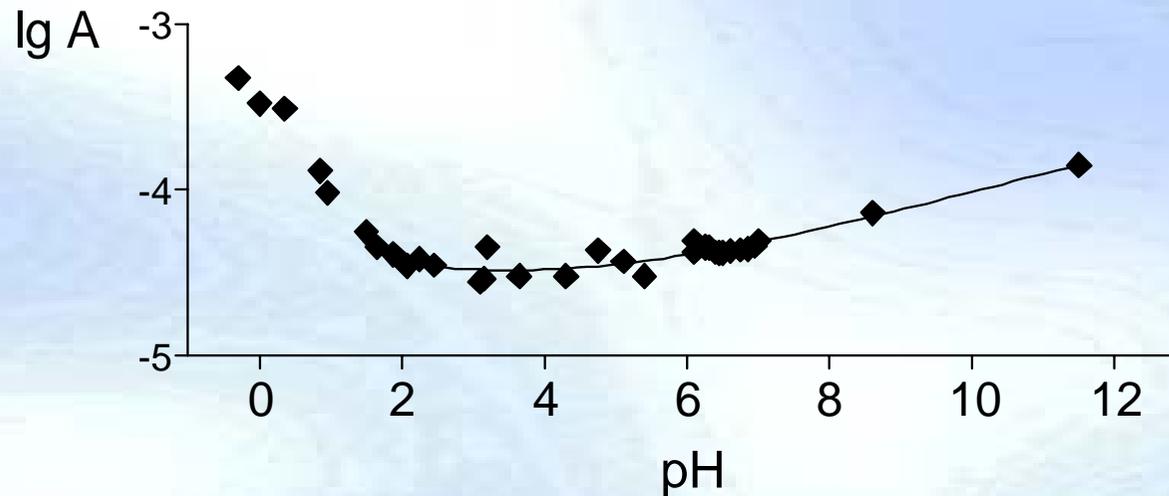
From the array of the experimental data (N points) the g-th point (only one) is excluded, the unknown parameters are calculated and the weighed residues of cross-validation are found:

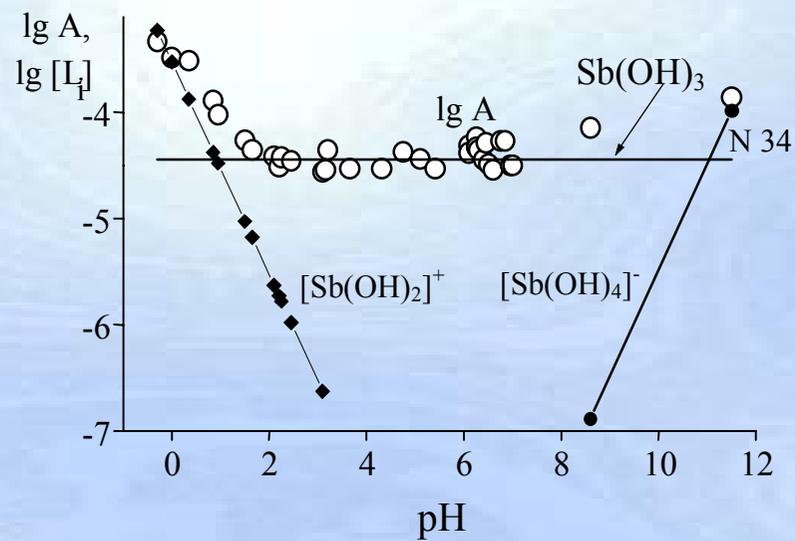
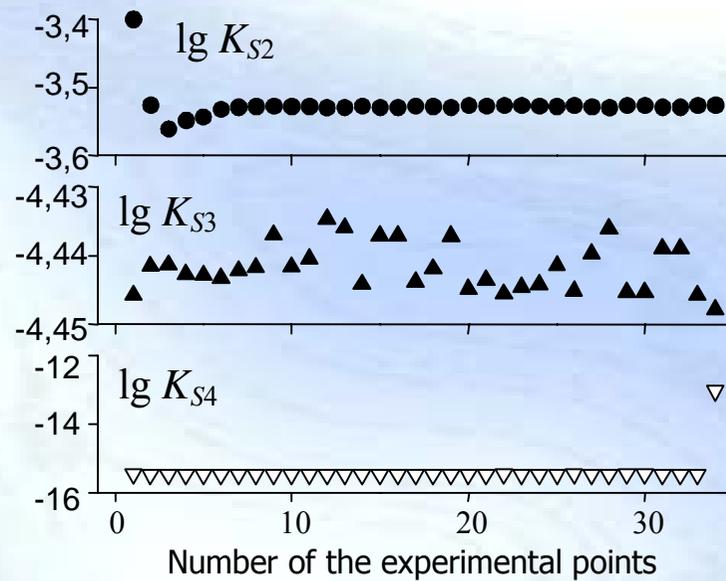
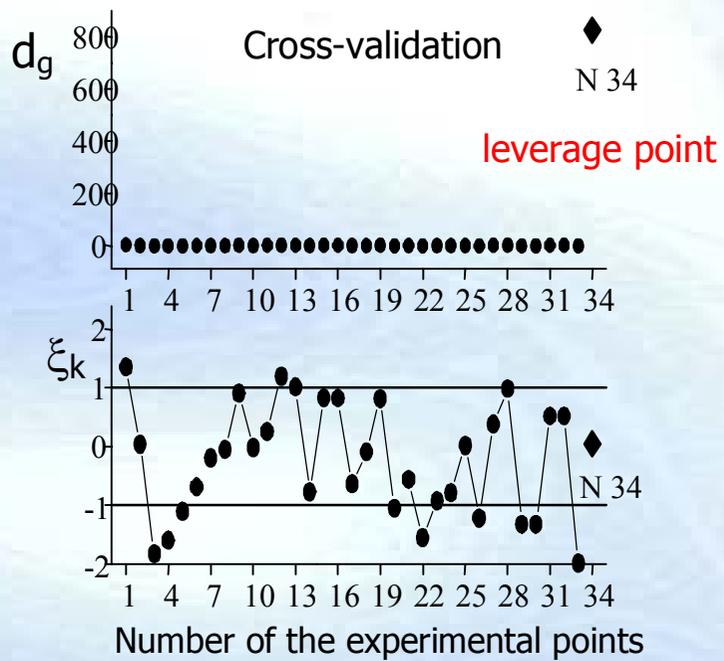
$$d_g = w_g^{1/2} \cdot (\hat{A}_g - A_g).$$

This procedure is repeated for all points and the cross-validation residual variance is

calculated:  $\tilde{s}_0^2 = \sum_{g=1}^N d_g^2$ .

## The hydrolytic equilibria of Sb(III) (from the data on the $\text{Sb(OH)}_3$ solubility)





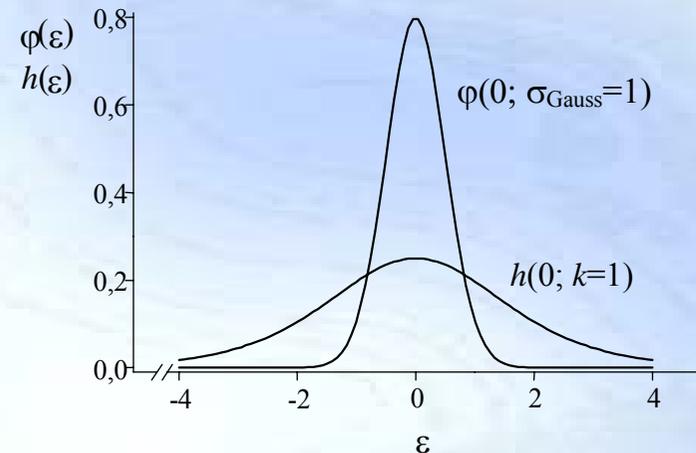
## The Huber's M-estimates as the alternative to the least squares method

$$\rho(\varepsilon) = [(100 - \delta) \cdot \varphi(0, \sigma_{\text{Gauss}}) + \delta \cdot h(\varepsilon)] / 100$$

$\varphi(0, \sigma_{\text{Gauss}})$  – the density of the Gauss distribution,

$h(\varepsilon)$  – the density of gross errors distribution (the distribution function with long tails),

$\delta$  – the percentage of gross errors, %.



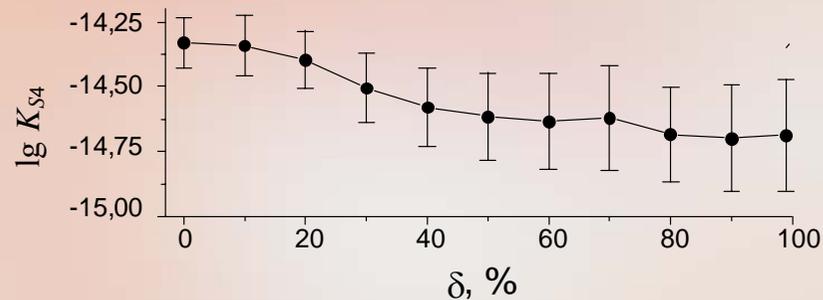
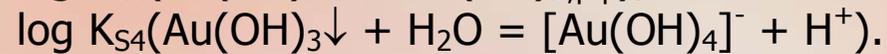
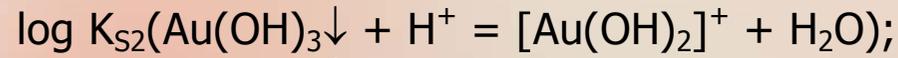
$$|\boldsymbol{\theta}\rangle = \arg \min \sum_{k=1}^N \rho(\xi_k; \boldsymbol{\theta}) \quad \xi_k = w_k^{1/2} \times (A_k^{\text{calculated}} - A_k^{\text{measured}})$$

$$\text{the loss function} \quad \rho(\xi) = \begin{cases} (1/2)\xi^2, & \text{at } |\xi| \leq c \\ c|\xi| - (1/2)c^2, & \text{at } |\xi| > c \end{cases}$$

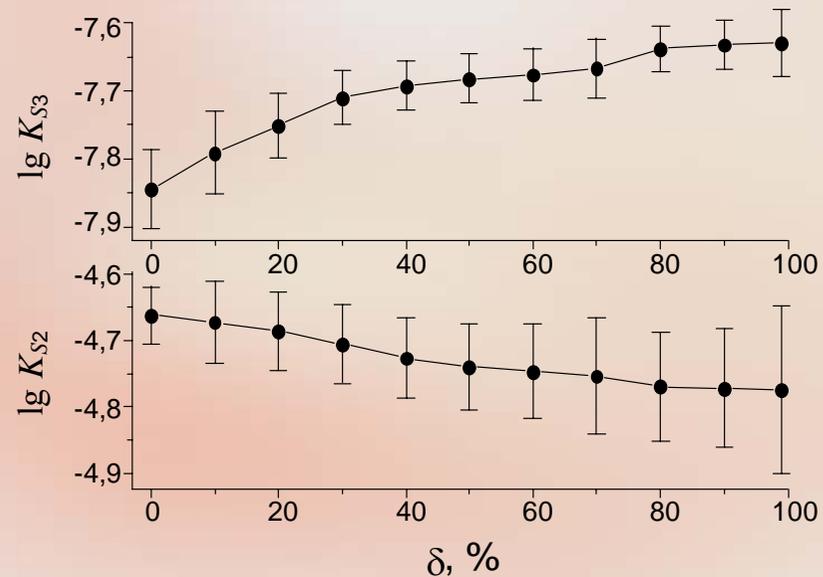
If  $\delta \rightarrow 0$  then  $c = \infty$  (the Huber's M-estimates are the same as the least-squares ones).

If  $\delta \rightarrow 100\%$  then  $c \rightarrow 0$  (the Huber's M-estimates are the same as the least-modules ones).

## Hydrolytic equilibria of Au(III) (from the data on Au(OH)<sub>3</sub> solubility)



Variation of the calculated  
M-estimates of log K<sub>Si</sub> on δ.



**The accepted values:**

$$\log K_{S2} = -4.77 \pm 0.08;$$

$$\log K_{S3} = -7.64 \pm 0.03;$$

$$\log K_{S4} = -14.68 \pm 0.15.$$

## CLINP 2.1: PROGRAM FOR ROBUST PARAMETRIC IDENTIFICATION OF COMPLEXATION MODELS

<http://www.bestnet.kharkov.ua/kholin/clinp.html>

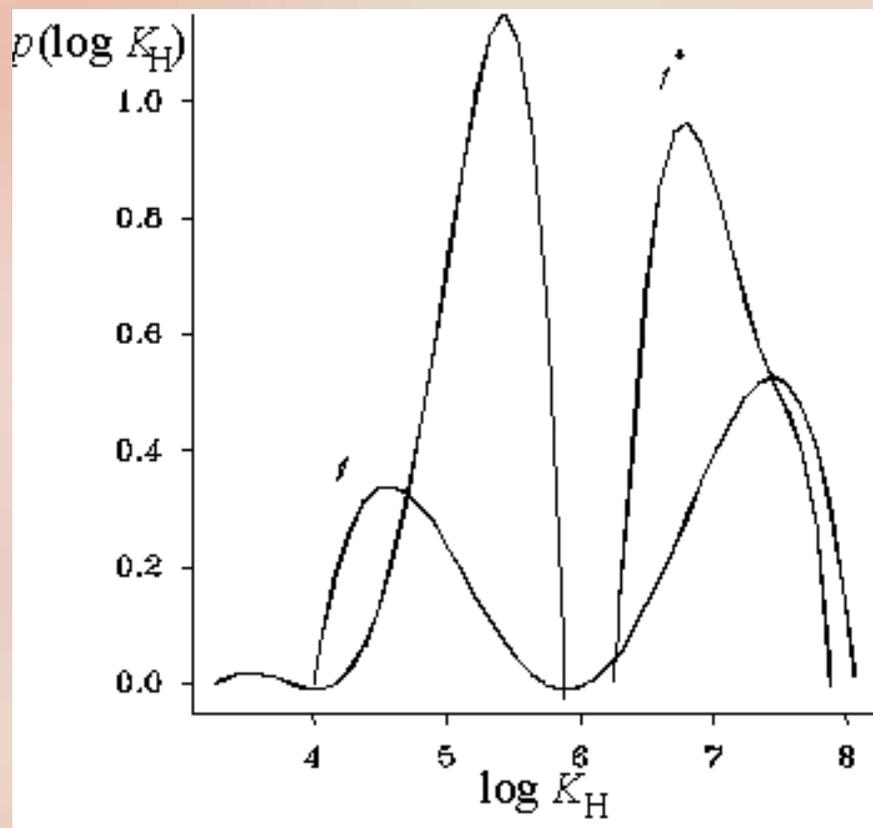
### APPROACH

- **NEW**: calculation of parameters on the base of robust Huber's M-estimates; insensitivity to outliers;
- **NEW**: modified globally convergent and rapid algorithms of Newton and Gauss-Newton methods;
- **NEW**: tools for convenient data analysis (grid calculation, cross-validation, SVD decomposition of Jacobi matrix, pointing to predominant components, etc.);
- **NEW**: adaptive system of data preparation for different experimental methods;
- easy to handle experimental data from different methods.

### RESULTS

- **logarithms of equilibrium constants** for reactions in solutions, extraction systems or sorption systems;
- **covariance matrix** and **correlation coefficients** (multiple, total and partial) for estimated parameters;
- **intensity factors** for complexes (e.g. molar absorptivities);
- **local criteria** of model adequacy – weighted residuals;
- **global criteria** of model adequacy (residual variance,  $\chi^2$ , skewness and kurtosis for distribution of residues, mean residual, residual mean);
- **NEW: cross-validation**;
- **NEW**: diagnostics of the **ill-posed** (redundant) chemical models;
- **equilibrium concentrations** of all reagents at all experimental points;
- **NEW: predominant components**.

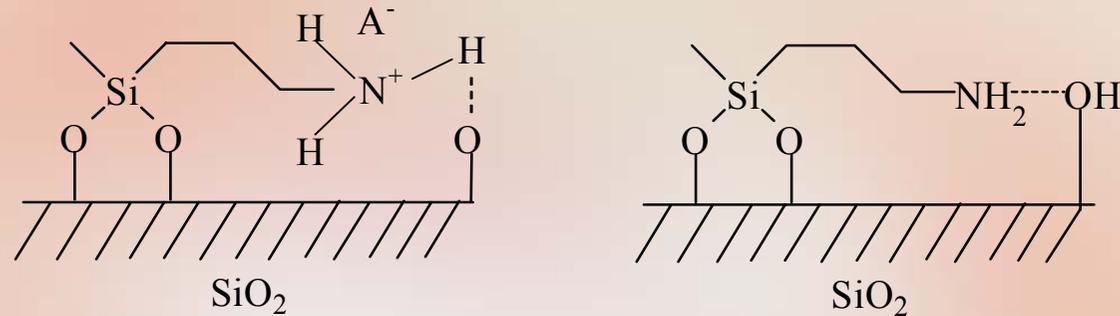
**Studying the energetic heterogeneity of 35 silica samples with aliphatic amines grafted on the surface**



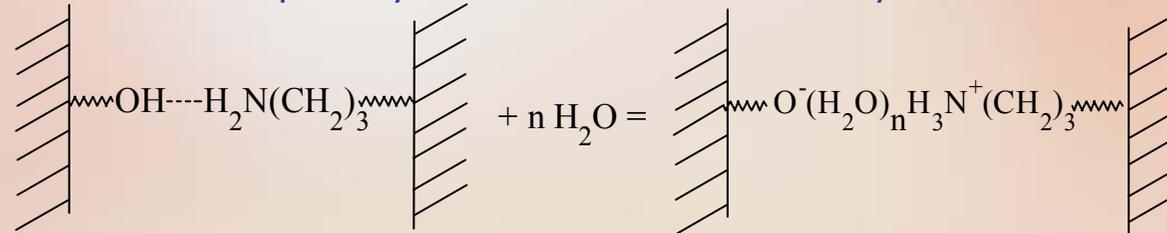
**The distribution functions  $p(\log K_H)$  for amines grafted on silica surface.**

**1\* corresponds to aminosilica 1 after "heating – cooling" cycle.**

The system of hydrogen bonds or salt-like products and arched structures at the surface:



Heating accelerates the complete hydration of the sub-surface layer:

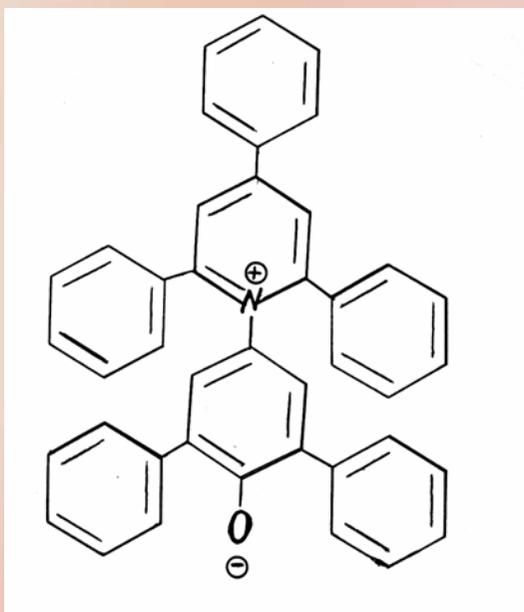


The full description was published in

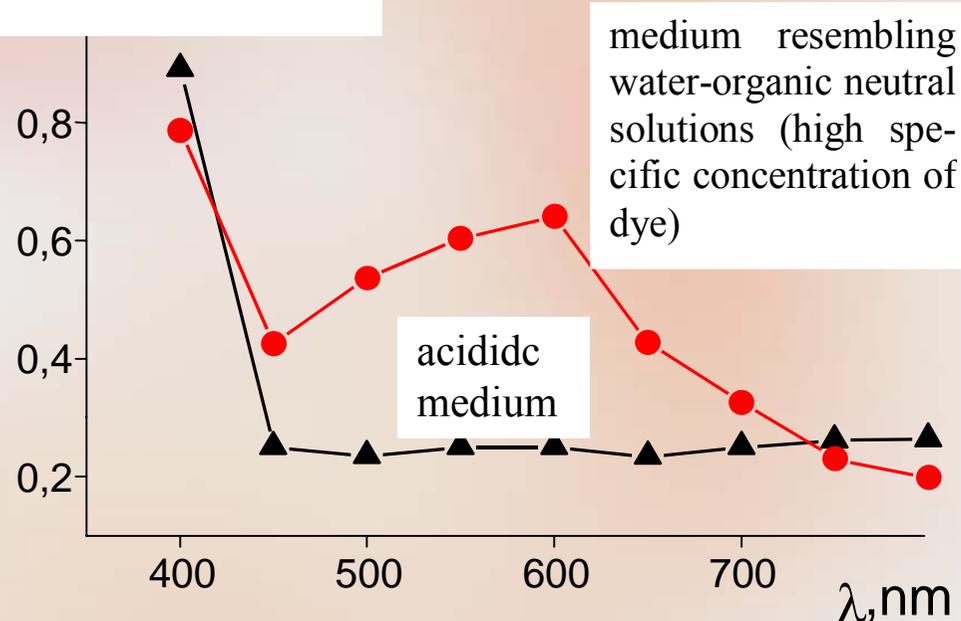
- a) **Kholin Yu., Myerniy S., Varshal G. *Adsorption Science & Technology*. 2000. 18, 267-294.**
- b) **Kholin Yu., Myerniy S., Shabaeva Y., e.a. *Adsorption Science & Technology*. 2003, 21, 53-66.**

The direct evidence of the unrandom surface topography of aminosilicas was obtained.

The absorbance spectra of solvatochromic indicator  $E_T^{30}$  adsorbed at the aminosilica surface



Kubelka-Munk function



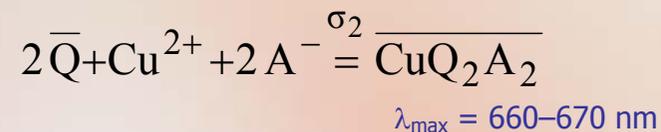
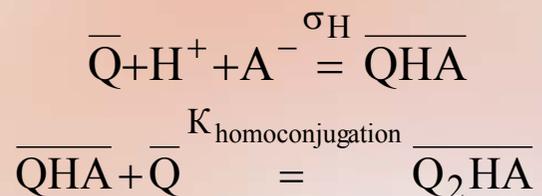
Khristenko I.V., Kholin Yu. V., Mchedlov-Petrosyan N.O.,  
Probing the surface of aminosilicas by solvatochromic  
betaine Reichardt's dye, Kharkov Univ. Bul., 2002. No 549.  
Chemical series. Issue 8 (31). 115-118.

## The QPCA as a tool to explore the thermodynamic features of functionalized silicas

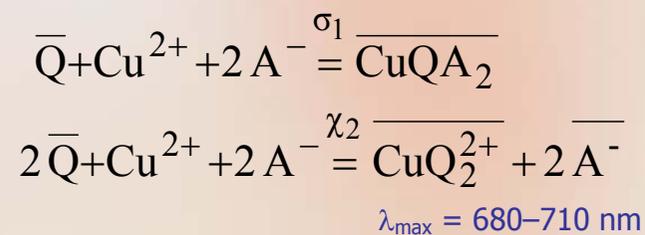
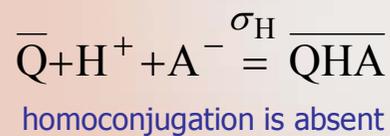
Model of protolytic equilibria

Model of complexation equilibria

(20 – 40) °C

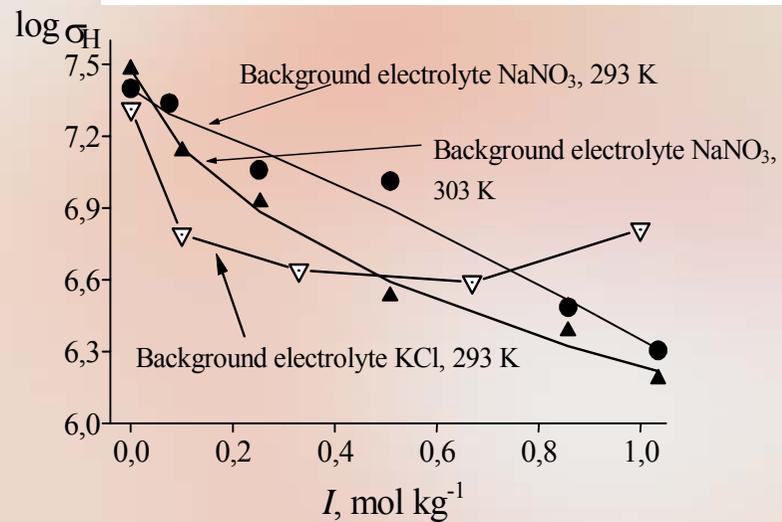


50 °C

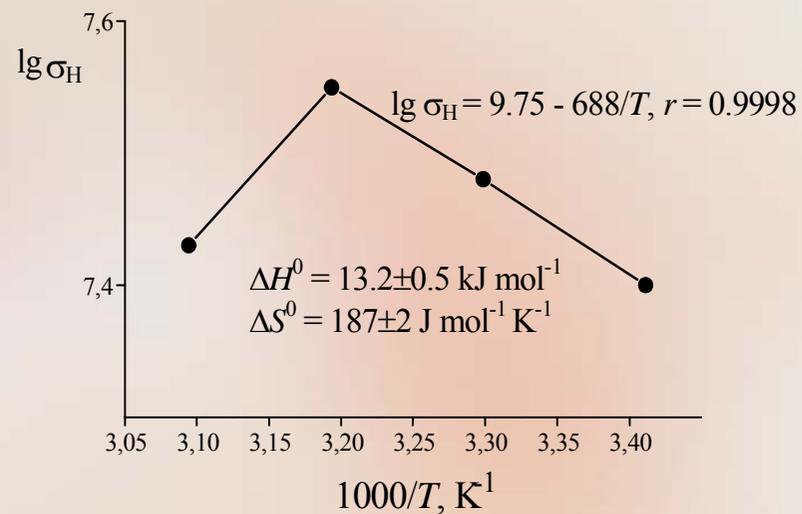


Q is organic amine, grafted on silica surface

Approximation on the base of the Pitzer method

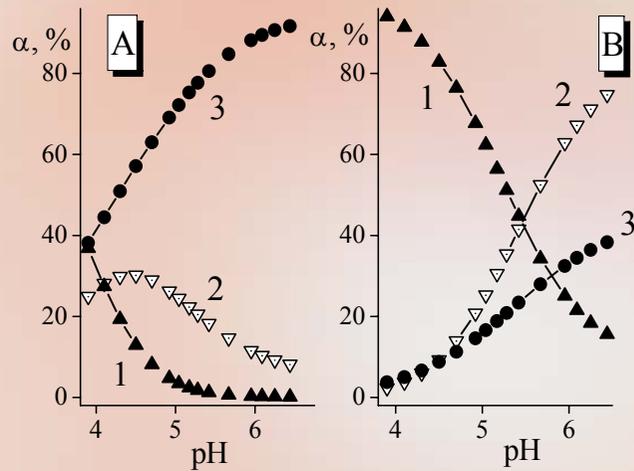


The temperature dependence of the thermodynamic protonization constants



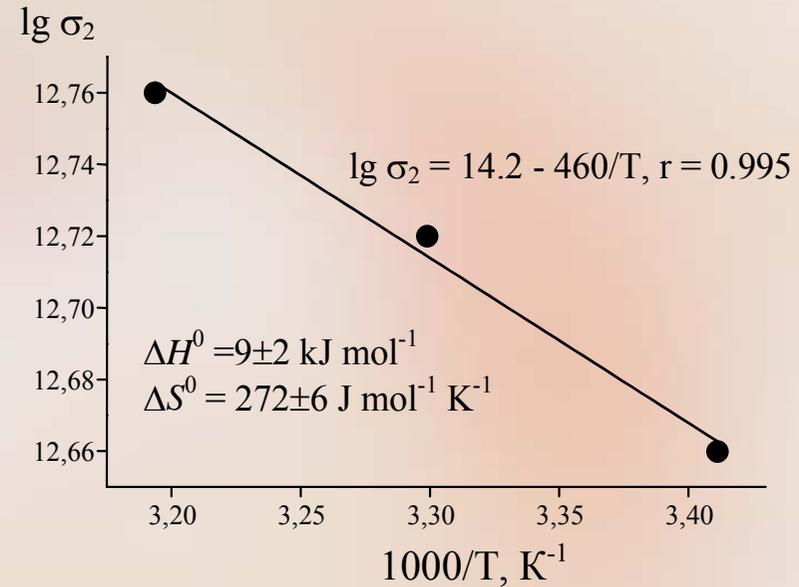
Marques H., *J. Chem. Soc. (Dalton)*, 1991:  
 n-propylamine:  $\Delta H^0_{298} = -57.4 \text{ kJ mol}^{-1}$ ,  $\Delta S^0_{298} = 17 \text{ J mol}^{-1} \text{ K}^{-1}$ ,  
 ammonia:  $\Delta H^0_{298} = -54.1 \text{ kJ mol}^{-1}$ ,  $\Delta S^0_{298} = \sim 0$ .

### The temperature dependence of the thermodynamic complexation constants



Degrees of formation of copper complexes. 50 °C,  
 $I = 0.25 \text{ mol L}^{-1}$  (A) and  $0.75 \text{ mol L}^{-1}$  (B).  $c^0(\text{Cu(II)}) / c^0(\text{amine}) = 0.14$

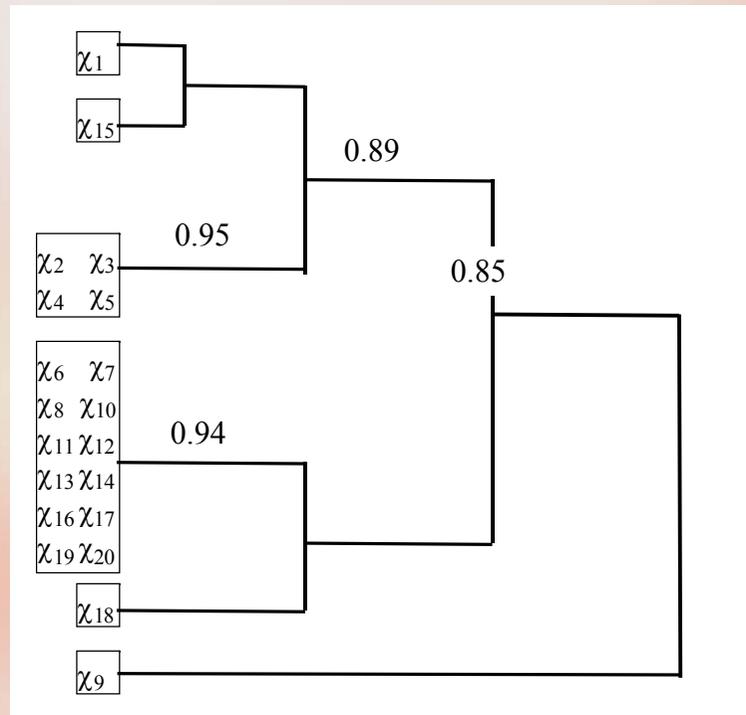
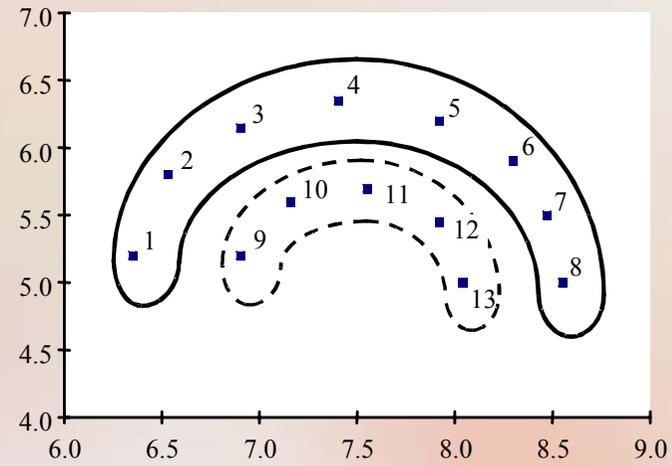
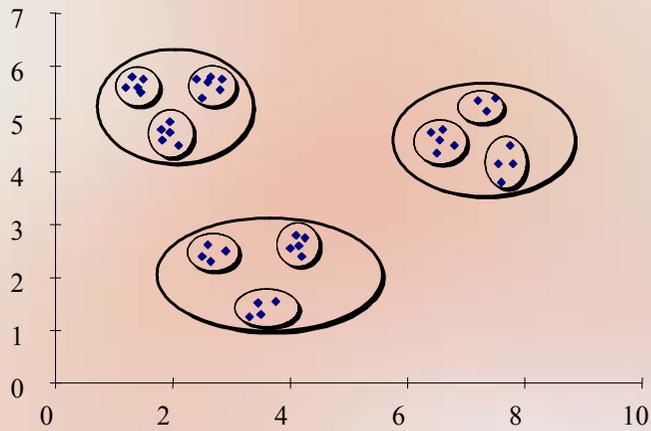
1 –  $\text{Cu}^{2+}$ , 2 –  $\text{CuQA}_2$ , 3 –  $\text{CuQ}_2^{2+}$ .



**For  $\text{NH}_3$ :**

$\Delta H^0_{298} = -43.6 \text{ kJ mol}^{-1}$ ,  $\Delta S^0_{298} = -0.2 \text{ J mol}^{-1}\text{K}^{-1}$

## Fuzzy classification of the incomplete arrays of data about physico-chemical properties of functionalized materials



## Another substances and materials studied by the QPCA method

### *In cooperation with*

•Humic and fulvic acids and coals (protolytic properties, complexation, sorption of heavy and noble metals)

Prof Galyna Varshal and Dr Irina Koscheeva (V.I.Vernadskii Institute of Geochemistry and Analytical Chemistry, Moscow)

•Silica-organics xerogels (protolytic and complexation properties)

Prof Edilson Benvenuto and Prof Tania Kosta (University of Porto Alegre, Brazil)

•Cellulose acetate/ $\text{Al}_2\text{O}_3$ /inorganic or organic ligand hybrid materials (sorption and ion exchange)

Prof Yoshitaka Gushikem (University of Campinas, Brazil)

•Non-covalently immobilized dyes and compleximetric indicators (adsorption equilibria, complexation, surface probing)

Prof Nykolay Mchedlov-Petrosyan (Kharkov University) and Prof Christian Reichardt (Philipps University, Marburg)

## Research team

- Senior Research Fellow  
Dr Sergiy Myerniy
- Research Fellow Dr  
Dmitry Konyaev
- Junior Research Fellow  
Mrs Inna Khristenko
- Engineer Mrs Olena  
Shulichenko
- Research Fellow Mr  
Sergiy Korneev
- Post-graduate student  
Mrs utta Shabaeva
- Graduate Student Mr  
Anton Panteleimonov



# My city Kharkov



