УНІВЕРСИТЕТ - 200   *ХІМІЧНИЙ ФАКУЛЬТЕТ - 110*   *НДІ ХІМІЇ - 75*

--*Prōspectē* --

ENERGETIC HETEROGENEITY OF SORBENTS:
NUMERICAL CALCULATION OF AFFINITY DISTRIBUTIONS
© 2004 Yu. Kholin, S. Myerniy

Холин Юрий Валентинович. Химик, доктор химических наук, профессор. Докторская диссертация: «Количественный физико-химический анализ комплексообразования в растворах и на поверхности химически модифицированных кремнеземов: содержательные модели, математические методы и их приложения» (2000 год). Родился в 1962 году в Харькове. В 1984 году окончил химический факультет Харьковского государственного университета, в 1984-1987 годах обучался в аспирантуре, в 1987-1991 годах работал старшим научным сотрудником НИИ химии, с 1991 года – доцент химического факультета, с 2001 года – заведующий кафедрой химического материаловедения, с 2004 года – проректор по учебно-методической работе. Создал научное направление – количественный физико-химический анализ комплексообразования на поверхности функционализированных гибридных материалов, в рамках которого получены приоритетные результаты в химии функционализированных кремнеземов.Основные научные интересы: физическая и координационная химия поверхностно модифицированных гибридных материалов, применение гибридных материалов в анализе, хемометрия и теория анализа данных, химическое образование.

Мерный Сергей Алексеевич. Химик, кандидат химических наук. Кандидатская диссертация: «Разработка, верификация и применение моделей, описывающих равновесия сорбции компонентов растворов комплексообразующими кремнеземами» (1997 год). Родился в 1968 году в Харькове, в 1992 году закончил химический факультет Харьковского государственного университета, с этого времени работает в университете младшим научным сотрудником, научным сотрудником, старшим научным сотрудником. Предложил новые методы решения некорректных задач, возникающих при описании энергетической неоднородности сорбентов, внедрил в практику количественного физико-химического анализа робастные алгоритмы на основе М-оценок Хьюбера, разработал программное обеспечение для моделирования равновесных систем. Основные научные интересы: хемометрия, математически некорректные задачи физической химии, применение теории информации в химии, робастное оценивание в теории анализа данных.

THE PHENOMENON OF ENERGETIC HETEROGENEITY AND ITS QUANTITATIVE CHARACTERISTICS

REVIEW OF COMPUTATIONAL METHODS USED TO CHARACTERIZE THE ENERGETIC HETEROGENEITY

ALGORITHM CAS (<u>C</u>OMPUTED <u>A</u>FFINITY <u>S</u>PECTRUM)

ALGORITHM BASED ON THE MAXIMUM ENTROPY APPROACH

Summary. The numerical characterization of energetic heterogeneity of sorbents is discussed. This includes the ill-posed problem of solution of the first kind integral equations that leads to great computational difficulties. The computational methods used in this field are compared and the ways to provide reliable solutions are marked. The proposed non-iterative and rapid algorithm CAS allows to reduce the possible set of solutions to a compact and applies the Tikhonov α-regularization and the cross-validation techniques for enhancing the numerical stability of calculations. Also, a novel method based on the maxent approach has been developed as an alternative to the traditional procedures. The method was shown to be capable of generating affinity distributions close to model ones even in the case of narrow density functions and noisy experimental data.

The phenomenon of energetic heterogeneity and its quantitative characteristics

Physi- and chemisorption from solutions on mineral (hydr)oxides and chemically modified silicas, binding chemical species to humic acids and another amorphous solid surfaces [1-3] are often influenced by effects of energetic heterogeneity. In spite of the long history, the problem of estimating the energetic heterogeneity of materials is far from being resolved. Two types of heterogeneity should be distinguished. First of them, the biographic heterogeneity [4], is the intrinsic feature of materials. It may be due to the inhomogeneous topography of a surface, the presence of pores of different shape and size, the chemical heterogeneity of a surface, etc. The degree of occupation of binding sites by bonded species does not affect the scope of biographic heterogeneity. In the contrast, if the affinity of binding sites to fixed entity is changed in the course of sorption, this phenomenon is to be attributed to the evolutionary heterogeneity [5]. The most common causes of it are electrostatic interactions, for instance, changes of surface charge and electrostatic potential in dependence on the ion sorption degree.

Here we restrict ourselves to the simplest case of monocomponent binding, namely, fixation of species M from solution by binding centers $\overline{Q}$ * which leads to the formation of sorption complexes $\overline{MQ}$:

$$M + \overline{Q} \ \overset{K}{=} \ \overline{MQ} , \qquad (1)$$

where K is the affinity constant.

To characterize quantitatively the effects of energetic heterogeneity one should 1) to measure the dependence of $f([M])$ on $[M]$ where $0 \leq f([M]) \leq 1$ is a fraction of binding centers $\overline{Q}$ occupied with M, square brackets denote equilibrium concentrations of species in solution; 2) to postulate a model which allows to separate effects of biographic and evolutionary heterogeneities and 3) to calculate numerically the model parameters. To perform step 3 it is necessary to solve the linear first kind Fredholm equation:

$$f([M]) = \int_0^\infty \theta^{\text{local}}([M], K) \cdot p(K) \, dK , \qquad (2)$$

with respect to $p(K)$, where $p(K)$ is the nonnegative density function of affinity constants K; the kernel of the integral equation $\theta^{\text{local}}([M], K)$ is the local binding isotherm. The density function $p(K)$ describes the biographic heterogeneity, while the evolutionary heterogeneity is taken into consideration by choosing an appropriate expression to the local isotherm. Useful characteristics of density function $p(K)$ are

$$\text{mean value} \quad E = \int_0^\infty K \, p(K) \, dK , \qquad (3)$$

$$\text{variance} \quad D = \int_0^\infty (K - E)^2 \, p(K) \, dK , \qquad (4)$$

$$\text{skewness} \quad A = \int_0^\infty (K - E)^3 \, p(K) \, dK . \qquad (5)$$

In addition to $p(K)$, one can use the integral distribution function

$$P(K) = \int_0^K p(x) \, dx , \ \ 0 \leq P(K) \leq 1, \qquad (6)$$

to measure the biographic energetic heterogeneity. Up to date a lot of models have been proposed to describe the evolutionary heterogeneity in terms of electrostatic interactions (see, for instance, ref. [6-15]). It should be particularly emphasized that the shape of binding isotherm does not allow to

---

* Binding center $\overline{Q}$ is assumed to be an adsorption site on the surface of a solid sorbent, a ligand grafted on silica surface, a functional group of humic acid and so on.

decide what type of energetic heterogeneity or both of them take place [2]. Thus, there is no serious experimental guideline to prefer one or another function for a local isotherm. In the limiting case when both electrostatic and lateral interactions are negligible, the kernel $\theta^{local}([M], K)$ reduces to the Langmuir isotherm

$$\theta^{local}([M], K) = \frac{K \cdot [M]}{1 + K \cdot [M]} \, . \tag{7}$$

For the sake of computational simplicity, expression (7) has been adopted for the local isotherm in our work. It is possible to validate this decision *a posteriori*. For this, it is necessary to measure binding isotherms varying the ionic strength of solution (*I*), to calculate corresponding density functions $p(K)$ and to compare them. If the shapes of functions $p(K)$ and positions of maxima do not depend significantly on *I*, it may be concluded that the choice of the Langmuir isotherm is justified and effects of the evolutionary heterogeneity are negligible [16]. In the opposite case, the calculated functions $p(K)$ should be considered as the density functions for apparent affinity constants only. An additional advantage of using equation (7) for the local isotherm θ^{local} is a simple way to express the integral distribution function $P(K)$ in terms of primary experimental data:

$$\tau([M]) = 1 - f([M]) = \int_0^\infty P(K) \frac{[M]}{(1 + K \cdot [M])^2} \, dK \, . \tag{8}$$

To find exact solutions of equations (2) and (8) one has to know the analytical expression for $f([M])$ and to perform the integral transformations (such as Fourier, Laplace or Stiltjes). Then, if the local isotherm is specified by the Langmuir equation, the sought density and integral distribution functions may be represented as follows [17-19]:

$$p(x) = \frac{\partial f(x)}{\partial x} - \frac{\pi^2}{3!} \frac{\partial^3 f(x)}{\partial x^3} + \frac{\pi^4}{5!} \frac{\partial^5 f(x)}{\partial x^5} - \frac{\pi^6}{7!} \frac{\partial^7 (x)}{\partial x^7} \pm \dots |_{x = \ln K = -\ln[M]} \, , \tag{9}$$

$$P(x) = \tau(x) - \frac{\pi^2}{3!} \times \frac{\partial^2 \tau(x)}{\partial x^2} + \frac{\pi^4}{5!} \times \frac{\partial^4 \tau(x)}{\partial x^4} - \frac{\pi^6}{7!} \times \frac{\partial^6 \tau(x)}{\partial x^6} \pm \dots |_{x = \ln K = -\ln[M]} \, . \tag{10}$$

Unfortunately, experimental function $f([M])$ is a table of values rather than an analytical expression. Hence, only approximate numerical methods are applicable to find distribution functions.

Review of computational methods used to characterize the energetic heterogeneity

As equations (2) and (8) are the first kind Fredholm integral equations, we deal with the ill-posed problem. Main computational difficulties are connected with this.

In particular, there exist many different possible solutions $p(K)$ which fit the measured $f([M])$ values within their experimental errors. Moreover, small fluctuations in primary experimental data ([M] or $f([M])$) may cause an arbitrary large fluctuations in the calculated density function $p(K)$. Also, oscillatory solutions with no physical meaning may easily arise. Consequently, all methods for practical use have to reject non-realistic (negative or oscillating) functions $p(K)$.

The approximate methods for solving equation (2) may be divided into two groups.

To avoid oscillating solutions $p(K)$ within the methods of "local isotherm approximation", the ill-posed problem is transferred to a better posed one through changing the continuous local isotherm with an approximate discontinuous (stepwise) function. Zeldowitch [20] and Roginskii [21] were first to employ this approach. The Roginskii's method (known also as "condensation approximation") was comprehensively discussed and tested [1]; till now it is the most popular and commonly used tool for solving equation (2). During last decades the methods of "local isotherm approximation" were further developed (see, for instance, LINA, LOGA-1, LOGA-AS approximations described by Nederlof et al. [22, 23]). Errors resulting from using these methods have been comprehensively studied. Though the novel approaches work better than the initial Roginskii's method [1], narrow distributions $p(K)$ are always flattened and sometimes the positions of maxima are detected incorrectly [23]. Also, the use of these methods for describing chemisorption equilibria is unfavorable from the theoretical point of view. Actually, if the chemical binding without lateral interactions takes place (in particular, when hydrogen or metal ions react with ligands grafted on silica surface or with functional groups of humic

substances) there exists only one expression the form of which is analogous to the mass-action law. This is the Langmuir equation.

The second group includes a wealth of general computer-oriented numerical methods. Their employment consists in the search of a suitable balance between the goodness of fitting measured f values and the smoothness of solutions $p(K)$. The quality of approximation is commonly measured by the criterion

$$U = \sum_{k=1}^{N} \left(\hat{f}_k - f_k \right)^2 , \tag{11}$$

where k is the number of the experimental points; N is the total number of measurements; $\hat{f}_k$ is the f value calculated for the k-th point. Solutions $p(K)$ are found via minimization of criterion (11)[*], while the degree of smoothing of sought for functions $p(K)$ is determined by peculiarities of computational procedures.

Sometimes, the total binding isotherm $f([M])$ is approximated by a smooth function $\tilde{f}([M])$ with few fitting parameters and the latter is substituted into equation (9) instead of $f([M])$. A lot of expressions (for instance, low-power polynomials, hyperbolas, cubic splines, the generalized Freundlich, Rothemund-Kernfeld, Jowanowic-Freundlich, Toth and UNILAN isotherms, the Jaroniec exponential equation) has been tried [24-28]. One of the most successful approaches was the approximation of experimental $f([M])$ dependence with a multi-Langmuir isotherm (as a rule, with the bi-Langmuir one) [29, 30]. Two disadvantages of this approach are easily seen. First, the searched solution $p(K)$ depends highly on the choice of function $\tilde{f}([M])$ while the latter seems to be completely arbitrary. Second, functions $\tilde{f}([M])$ have only one or two, seldom three or four fitting parameters. As a result, sometimes measured $f([M])$ values are approximated roughly, and function $p(K)$ is reproduced inauthentically.

The famous early algorithm CAEDMON [31, 32] approximated the integral in formula (2) by the sum calculated for equidistant knots K_j. For each experimental point equation

$$\hat{f}_k = \sum_{j=1}^{J} p(K_j) \frac{K_j \cdot [M]_k}{1 + K_j \cdot [M]_k}, \, k = 1, 2, \ldots, N, \tag{12}$$

was written. Here J is the number of knots inside the interval of variation of K, K_j is the value of K at the j-th knot, $p(K_j)$ is the fraction of binding centers with affinity constant K_j. So, the sought for function $p(K)$ was found as a superposition of Dirac δ-functions. The non-negative least squares (NNLS) algorithm [33] was applied to minimize (11) in respect to $p(K_j)$. Natural restrictions

$$\sum_{j=1}^{J} p(K_j) = 1, \, p(K_j) \geq 0, \tag{13}$$

did not prevent a numerical instability of computations at high J values.

General methods for solving first kind integral equations [34-36] came into use in this field. They were based on the approximation of integral $\int_0^{\infty} \theta^{local}([M], K) \cdot p(K) \, dK$ with quadratures. The algorithm CAESAR [1] as well as algorithms proposed by Leunberger and Schindler [37], Bratskaya and Golikov [38], Garmash $et\,al$. [39] have approximated the integral with expression

$$\hat{f}_k = \sum_{j=1}^{J} \frac{K_j \cdot [M]_k}{1 + K_j \cdot [M]_k} \cdot p(K_j) \Delta K, \quad J < N, \quad k = 1, 2, \ldots, N , \tag{14}$$

where K_j are equidistant within the interval $[K_{min}, K_{max}]$. To find unknown $p(K_j)$, the NNLS algorithm was applied. To fit measured $f([M])$ values within experimental errors it is necessary to decrease the

---

[*] Reasons to choose the sum of residual squares as the criterion to be minimized and possible alternatives have not been discussed in the literature. Perhaps, the main reason to give preference to the least squares estimates was connected with the desire to simplify computations to the utmost.

integration step ΔK. Unfortunately, if ΔK is small the solution $p(K)$ is obtained as a "saw" with big teeth. In some cases the numerical stability of calculations was increased with the aid of the singular values decomposition (SVD) [33] included into the NNLS algorithm [1, 37]. Using the SVD technique has to be considered as a version of the α-regularization method (see below). Also, the discretization of the integral was performed in the EM algorithm [40, 41], though the authors have used their own iterative procedure instead of NNLS to compute unknown $p(K_j)$ values. It has been shown that the EM algorithm converges and never leads to negative $p(K_j)$. Moreover, the proper guess of $p(K_j)$ prevents oscillating solutions. The main disadvantage of this algorithm is a great number of iterations needed for convergence (up to 200 000).

To make the problem a better posed one, some authors [26, 30, 42, 43] searched a solution for $p(K)$ in the form of an analytical function with few fitting parameters. In particular, $p(K)$ was presented as a linear combination

$$p(K) = \sum_{i=1}^{X} b_i \cdot \varphi_i(K), \tag{15}$$

where $\varphi_i(K)$ are the Gaussian probabilities or γ-functions. Their parameters and coefficients b_i were determined by the minimization of criterion (11). In some cases this heuristic approach does not ensure accurate estimating $p(K)$ because at small X values there is no guarantee that the chosen approximation (15) is capable of approximating measured $f([M])$ values within experimental errors, while increasing X may cause the numerical instability of calculations. Besides, if there are many parameters to be determined, several local minima of criterion (11) may exist, and minimization loses unambiguity.

To handle properly the ill-posed problem and to obtain the numerically stable solutions, an additional *a priori* information about the searched functions should be involved. The way to use such information was developed by Tikhonov who proposed the general methods for solving ill-posed problems [44]. Also, he introduced a concept of conventionally correct problems [45]. To reduce an ill-posed problem to a conventionally correct one it is necessary to have a quantitative information which allows to choose one solution from many possible solutions. For instance, the problem turns into conventionally correct one if the information is available which restricts a set of possible solutions to a compact. An example of the compact is a set of non-decreasing functions with upper and lower bounds. The solution of a conventionally correct problem on the compact is numerically stable.

If the desired quantitative information is not available, the problem becomes an essentially incorrect one [45]. In this case, to ensure the numerical stability of calculations it is necessary to assume the smoothness of searched solution. Within the mentioned above numerical algorithms this is done implicitly. In our mind, the method of Tikhonov's α-regularization is of more objective nature. Applying the method, one should minimize, in $p(K)$, the new functional

$$U_\alpha = U + \alpha \cdot \Omega(p). \tag{16}$$

rather than criterion (11). Here $\Omega(p)$ is a nonnegative dampening function of $p(K)$, α is a positive regularization parameter. As noted above, conditions (13) have to be imposed on the searched function. House was the first who used the α-regularization method in estimating the energetic heterogeneity [46].

In the general case of the n-order α-regularization the regularizer $\Omega(p)$ is specified as follows [47]:

$$\Omega(p) = \int_0^\infty \sum_{j=0}^{n} q_j \cdot \left[\frac{d^j p(K)}{d K^j} \right]^2 d K, \tag{17}$$

where all coefficients $q_j \geq 0$ while $q_n > 0$. As a rule, more simple expressions are used, for instance:

$$\Omega(p) = \int_0^\infty \left\{ p^2(K) + n \cdot \left[\frac{d p(K)}{d K} \right]^2 \right\} d K, n \geq 0, \tag{18}$$

$$\Omega(p) = \| p(K) \| = \int_0^\infty p^2(K) \, d K. \tag{19}$$

Minimization of criterion (16) ensures a desired trade-off between the goodness of fit (U is a small parameter) and the smoothness of the solution ($\Omega(p)$ is a small parameter). With decreasing α, both the quality of $f([M])$ approximation and the value of dampening function $\Omega(p)$ increase whereas the smoothness of $p(K)$ and numerical stability of calculations fall off. So, the attainment of a suitable balance is ground on the correct choice of α. There are two general ways to specify α value. The method of "generalized discrepancy" requires the comprehensive information about the errors of primary experimental data [45, 47]. Usually computations of $p(K)$ are ambiguous because of the lack of this information. A more recent statistical method of cross-validation is aimed to overcome this demerit.

According to [48, 49], the basic idea of the cross-validation is to estimate the predictive rather than the fitting ability of the model. To make the cross-validatory choice of the regularization parameter from the set of admissible values $\alpha_{min} \leq \alpha \leq \alpha_{max}$, the following steps are required. 1) To exclude one observation with the number g from the original data set containing N experimental points; 2) to estimate unknown parameters by minimizing criterion (16) taking into account restrictions (13); 3) to find the g-th cross-validatory discrepancy:

$$d_g = \hat{f}_g - f_g \, , \qquad (20)$$

where $\hat{f}_g$ is computed using eq. (2) with the found density function $p(K)$ and measured $[M]_g$; 4) to repeat steps $1 - 3$ for all experimental points and to calculate the cross-validatory sum of residual squares

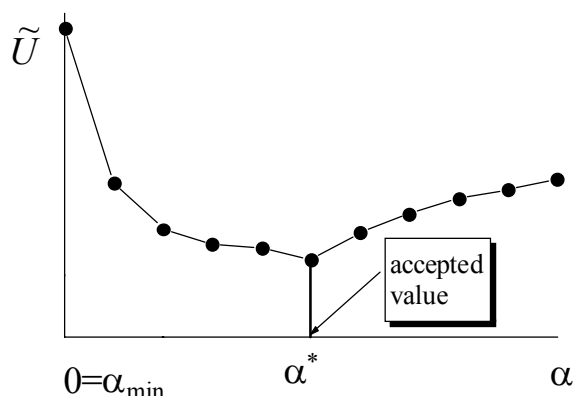$$\tilde{U} = \sum_{g=1}^{N} d_g^2 \, . \qquad (21)$$

Fig. 1. A possible dependence of the cross-validation criterion on the regularization parameter α.

Fig. 1 shows the expected variation of $\tilde{U}$ with α. At small α function $\tilde{U}(\alpha)$ has relatively large values because the experimental noise affects strongly the calculated functions $p(K)$ which are very far from being smooth. With increasing α function $\tilde{U}(\alpha)$ decreases, reaches a minimum and then increases again since solutions $p(K)$ found at high α are far from the exact one and the model is inadequate to describe the data. Hence, $\tilde{U}(\alpha)$ must possess at least one minimum. Having found the position of the first minimum, one can properly choose the value of the regularization parameter α. Unfortunately, this fairly objective and clear method is becoming time consuming with the increase of the number of experimental points N.

The strategy presented above was never been used before for the characterization of the energetic heterogeneity. At the same time, it was successfully implemented for solving a similar problem, namely, the analysis of the first-order rate constant spectra [50].

The encouraging results were obtained by Nederlof and co-authors [22] who employed the Tikhonov regularization and the "generalized cross-validation" (a less time-consuming version of

cross-validation [51, 52]) to find a quintic spline $\tilde{f}([M])$ which fits properly primary experimental dependence $f[M]$. The spline was searched via minimizing criterion

$$\frac{1}{N}\sum_{k=1}^{N}\left\{\tilde{f}_k([M]) - f_k([M])\right\}^2 + \alpha \int_{[M_1]}^{[M_N]}\left\{\frac{d^n\tilde{f}[M]}{d[M]^n}\right\}^2 d[M], \tag{22}$$

where $\displaystyle\int_{[M_1]}^{[M_N]}\left\{\frac{d^n\tilde{f}[M]}{d[M]^n}\right\}^2 d[M]$ is the smoothing regularizer. Instead of conditions (13), another

restrictions following from the physical meaning of the problem were imposed on the searched spline function:

$$\frac{dt(M)}{d[M]} \geq 1, \quad \frac{d^2 t(M)}{d[M]^2} \leq 0, \tag{23}$$

where $t(M)$ is the total (analytical) concentration of the species M to be bond. Calculated spline $\tilde{f}([M])$ was used to find the distribution function $p(K)$ with the aid of methods of "local isotherm approximation". The technique discussed above have pretensions to calculate $p(K)$ without the full information about the data errors. But it should be noted, that the unique solution of an ill-posed problem can not be found if the data errors are unknown [53, 53].

The conclusion is that numerical methods of solving equation (2) may produce numerically stable and correct density functions $p(K)$ if the ill-posed (essentially incorrect) nature of the problem is taken into account; the cross-validatory version of the α-regularization method seems to be the reliable and promising tool.

Algorithm CAS (**C**omputed **A**ffinity **S**pectrum)

The following question may arise in this context. If only modern complicated methods may ensure numerically stable and fairly exact computation of density functions $p(K)$, why much more simple algorithms rising from the old Adamson and Ling graphical method [55] work well? To answer the question it is necessary to take into account that the Adamson-Ling' algorithm and its further versions, HILDA [46] and Quasi-Adamson [30], are based on the iterative calculations of the integral distribution functions $P(K)$ rather than the density ones. Estimating $p(K)$ does not affect the computational process. As integral distribution functions $P(K)$ belong to the compact set of restricted $(0 \leq P(K) \leq 1)$ and non-decreasing functions, computation of $P(K)$ on a compact is of the conventionally correct nature. This manifests itself in the numerical stability of calculations. Estimating density function $p(K)$ by differentiating the integral function $P(K)$ remains the ill-posed problem. For this reason the mentioned algorithms may generate spurious $p(K)$ peaks. Keeping in mind this interplay between the conventionally correct problem in the case of $P(K)$ and essentially incorrect problem in the case of $p(K)$, we have suggested our own algorithm for estimating the biographic energetic heterogeneity (its early versions were described elsewhere [56, 57]).

The use of expression (10) makes it possible to avoid the iterative procedure in calculating $P(K)$. This requires to calculate the derivatives of $\tau(x)$. As $\tau(x)$ are specified by a table of values, the ill-posed nature of the problem remains. To solve it we have to choose an approximation of measured $\tau(x)$ values. The simplest decision is to assume that $\tau(x)$ belongs to such analytical functions as polynomials, to expand $\tau(x)$ in a power-series about a point a within the x variation range, to keep first m terms of the expansion and to drop all the rest high-order terms:

$$\tau(x) = \sum_{i=0}^{m}\left(\frac{g_i(a)}{i!}\right)(x-a)^i, \tag{24}$$

where coefficients

$$g_i(a) = \left(\frac{d^i\tau}{dx^i}\right)_{|x=a}. \tag{25}$$

Substituting expression (24) into equation (10) gives relationship to calculate $P(x)$:

$$P(x) = \sum_{i=0}^{m} g_i(a)\, D_i(x)\,, \tag{26}$$

where $D_i(x)$ are the known functions of x:

$$D_i(x) = \begin{cases} \sum_{l=0}^{i/2} \dfrac{(-1)^l \pi^{(i-2l)}}{(i-2l+1)!} \cdot \dfrac{(x-a)^{2l}}{(2l)!}, & \text{for even } i, \\[2ex] \sum_{l=0}^{(i-1)/2} \dfrac{(-1)^{(l+1)} \pi^{(i-2l-1)}}{(i-2l)!} \cdot \dfrac{(x-a)^{(2l+1)}}{(2l+1)!}, & \text{for odd } i, \\[2ex] 1, & i = 0. \end{cases} \tag{27}$$

Hence, only numerical values of coefficients $g_i(a)$ are needed to calculate $P(x)$. To find them the system of linear algebraic equations is constructed:

$$\tau(-\ln[M]_j) = \sum_{i=0}^{m-1} g_i(a) \times \frac{(-\ln[M]_j - a)^i}{i!}, \quad j = 1, 2, ... N, \ N > m\,, \tag{28}$$

where j is the number of experimental point. Now it is possible to make use the information that $P(x)$ belongs to a compact set of non-descending $(p(x) = d\,P(x)\,/\,d\,x \geq 0)$ and restricted $(0 \leq P(x) \leq 1)$ functions. From equation (26) it is easy to obtain the expression for $p(x)$:

$$p(x) = \sum_{i=1}^{N-1} g_i(a) \cdot G_i(x)\,, \tag{29}$$

where $G_i(x) = dD_i(x)/dx$:

$$G_i(x) = \begin{cases} \sum_{l=0}^{i/2} \dfrac{(-1)^l \pi^{(i-2l)}}{(i-2l+1)!} \cdot c, & \text{for even } i, \\[2ex] \sum_{l=0}^{(i-1)/2} \dfrac{(-1)^{(l+1)} \pi^{(i-2l-1)}}{(i-2l)!} \cdot \dfrac{(2l+1)(x-a)^{(2l+1)}}{(2l+1)!}, & \text{for odd } i, \end{cases} \tag{30}$$

$$c = \begin{cases} 0, & l = 0, \\[2ex] \dfrac{2l\,(x-a)^{2l}}{(2l)!}, & l \geq 0 \end{cases} \tag{31}$$

Imposing the following restrictions on $P(\ln K)$ and $p(\ln K)$

$$0 \leq P_i(\ln K) = \sum_{i=0}^{m-1} g_i(a) \cdot D_i(-\ln[M]_j) \leq 1, \quad j = 1, 2, ..., N, \tag{32}$$

$$0 \leq p_i(\ln K) = \sum_{i=1}^{m-1} g_i(a) \cdot G_i(-\ln[M]_j), \quad j = 1, 2, ..., N, \tag{33}$$

we apply the SVD version of the NNLS algorithm [33] to solve the least squares problem (28) subject to linear inequality constraints (32), (33). Thus, the function $P(\ln K)$ is calculated on the compact without using an iterative procedure; simultaneously, the function $p(\ln K)$ is estimated without additional computations.

An important step of the algorithm is the proper choice of m value (the order of polynomial and the number of coefficients $g_i(a)$ to be determined). The cross-validatory search of m was included into the computational scheme.

Let us illustrate the features of the algorithm with the simulation example. For 22 pM^* values distributed randomly inside interval [0; 11] the $f(pM)$ values were calculated. The density function $p(\log K)$ was assumed to be the superposition of Gaussian and exponential probability functions:

$$p(\log K) = (p_1(\log K) + p_2(\log K))/2 , \qquad (34)$$

where

$$p_1(\log K) = \begin{cases} \exp(2-x), & \log K \geq 0, \\ 0, & \log K < 0, \end{cases} \qquad (35)$$

$$p_2(\log K) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\log K - 7)^2}{2}\right). \qquad (36)$$

The normally distributed errors with zero mean and standard deviations 0.05 were introduced into the exact f values (Fig. 2). Then the density function was reconstructed from the obtained "pseudo-experimental" data. In spite of the high level of the imposed errors, the calculated functions were close to the exact ones (Fig. 3). Only in a vicinity of the point $\log K = 2$ where the distribution functions are discontinuous, the reconstructed functions approximate the model ones roughly.
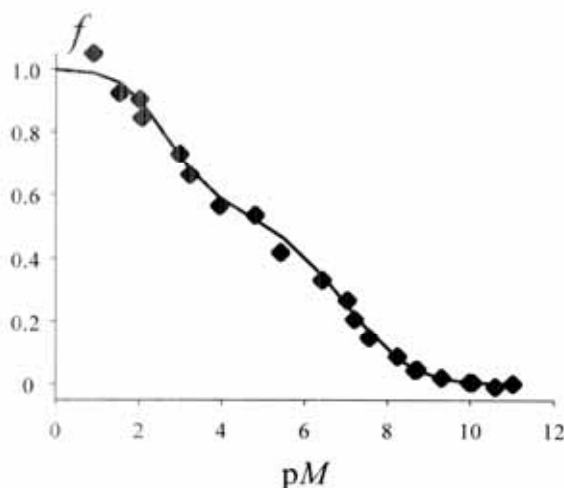
Fig. 2. Simulated binding degrees, f, versus pM. The solid line corresponds to the exact values, symbols ● denote data with introduced Gaussian errors.
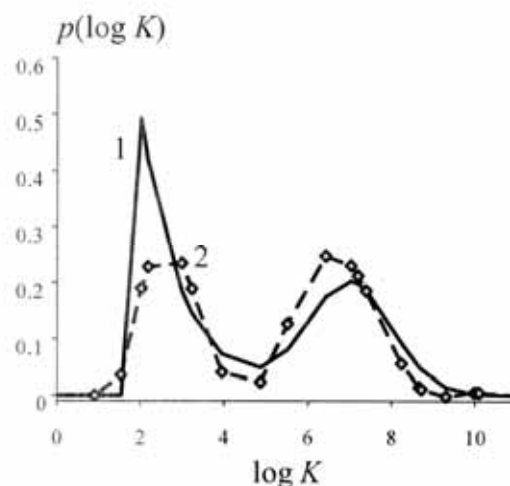
Fig. 3. The density functions calculated on the random grid with 22 knots. 1 – the model density function, 2 – the function reconstructed from data subjected to Gaussian random errors.

Algorithm based on the maximum entropy approach

The numerical methods for solving the inverse problems of the parametric identification can be divided into two groups: algebraic and probabilistic methods. The algebraic methods are based on the maximum likelihood principle and operate through the discretization of a priori models that are usually specified in the form of an integral equation, differential equation and so on. After that a certain criterion is minimized in respect to the searched discretizated solution (this criterion characterizes the quality of fitting i.e. the discrepancy between the experimental and the calculated data). The ill-posed nature of the problem manifests itself in the existence of many possible solutions compatible with the experimental data, and additional (external) information is necessary to select only one solution. The attraction of external information about searched solution is especially difficult if it is presented in the form of statistic or empirical rules. Also, the algebraic methods use the least squares method that, in accordance with the maximum likelihood principle, supposes the Gaussian distribution of experimental errors.

---

* $pM = -\log [M]$.

All methods discussed above belong to the algebraic ones. They require implicit or explicit assumptions about the smoothness of the searched affinity distribution $p(K)$ and/or the distribution law of the experimental errors ε_k. Due to the immanent properties of these methods, exploring the narrow density functions $p(K)$ is especially difficult.

The probabilistic methods assume that measured responses y_i are random quantities and search for the probability density $p_i(y_i)$. This group of methods includes, particularly, the Maximum Entropy (maxent) approach (the genesis of these methods in connection with the algebraic ones is excellently discussed in [58]). This approach can be considered as a peculiar ("probabilistic") form of regularization which is more natural for many experimental tasks.

In this work we suggest a novel, more unbiased method for calculating the affinity distributions that is based on the formalism of the maxent approach.

Remind that equation (2) may be rewritten without loss of generality in the discrete form (see equation (12)). Natural restrictions (13) follow from the physical sense of the problem.

The maxent discrete probability density function $p(K)$ is typically obtained by maximizing Shannon entropy [59]

$$S = -\sum_{j=1}^{J} p(K)_j \times \ln p(K)_j \ . \tag{37}$$

subjected to some known constraints:

$$\sum_{j=1}^{J} A_{kj}\, p(K)_j = Y_k \ , \ \ k = 1, 2, \ldots, N; \ \ N < J, \tag{38}$$

where A_{kj} are known coefficients and Y_k are values known from an experiment or theory [60, 61]. In other words, experimental data are considered as ***restrictions*** imposed on the searched density function $p(K)$. The routine variation problem is usually solved using the method of Lagrangian multipliers:

$$L(p,\lambda) = \int_K \left\{ \ln \frac{p(K)}{p_0(K)} + \sum_{k=1}^{N} \lambda_k (f_k - \hat{f}_k) \right\} p(K)\, dK \ , \tag{39}$$

where $p_0(K)$ is the *a priori* density function (as a rule, the uniform one). The solution in the stationary point $L(p,\lambda)$ is as follows:

$$p(K;\lambda) = \frac{1}{Z(\lambda)} p_0(K) \times \exp\left(\sum_{k=1}^{N} \lambda_k \hat{f}_k \right), \tag{40}$$

where

$$Z(\lambda) = \int_K p_0(K) \times \exp\left(\sum_{k=1}^{N} \lambda_k \hat{f}_k \right) dK \ , \tag{41}$$

with λ_k found from restrictions

$$\frac{\partial \ln Z(\lambda)}{\partial \lambda_k} = f_k, \ \ k = 1,\ldots, N \ . \tag{42}$$

In fact, the maxent approach allows to select the unique function $p(K)$ from many possible ones.

In order to use the maxent approach, we have transformed the set of equations (12) into the set of equations and inequalities imposed on the possible solutions $p(K_j)$:

$$\sum_{j=1}^{J} p(K_j) \times \frac{K_j \cdot [M]_k}{1 + K_j \cdot [M]_k} \geq f_k - \Delta \ , \ k = 1, 2, \ldots, N, \tag{43}$$

$$\sum_{j=1}^{J} p(K_j) \times \frac{K_j \cdot [M]_k}{1 + K_j \cdot [M]_k} \leq f_k - \Delta \ , \ k = 1, 2, \ldots, N, \tag{44}$$

$$\sum_{j=1}^{J} p(K_j) = 1, \ p(K_j) \geq 0, \tag{13}$$

where the threshold value Δ is the highest supposed value of ε_k.

To calculate $p(K_j)$, it is necessary to maximize entropy (37) with the account of restrictions (13), (43), (44). The minimum possible Δ value (Δ_{min}) is such a value that provides the compatibility of the system. Note, that this method does not involve any additional information on the experimental errors. If such information is available, it may be easily taken into consideration by the appropriate modification of inequalities sets (43) and (44). As usual, the computation of $p(K_j)$ is performed by the method of the Lagrangian multipliers.

Our method differs essentially from the original maxent approach. Namely, in our case the multiplicity of possible solutions is due to the uncertainty in Δ, while in the original method it follows from the incompleteness of restrictions.

It should be noted also that several algorithms using the maxent approach to solve numerically the Fredholm integral equations of the first kind have been developed [62, 63]. In these algorithms the Shannon entropy is used as the regularizer $\Omega(p)$ (see equation (16)). The capabilities of the entropy regularization methods are close to the Tikhonov α-regularization technique discussed above.

To test the capabilities of our method the following simulation was performed. It was assumed that three types of binding sites are present at the sorbent surface in equal concentrations, the corresponding $\log(K)$ values being equal to 3, 5 and 6. So,

$$p(\log K = 2) = p(\log K = 5) = p(\log K = 6) = \frac{1}{3}. \tag{45}$$

For 22 pM values distributed randomly within range [0; 11] the $f(pM)$ values were calculated. Then the random errors were introduced into the exact $f(pM)$ values, the relative errors having the Gaussian or uniform distributions with zero mean values and varied standard deviations. Two examples of the obtained "pseudo-experimental" data are presented in Fig. 4.

The density functions $p(K_j)$ were calculated with the aid of the algorithm described above and the corresponding S_{max} values, the values U of criterion (11) and the minimum possible threshold values Δ_{min} were found. The results of calculations are systematized in Tables 1 and 2. Examples of the reconstructed density functions are shown in Fig. 5 – Fig. 8. It is evident that the proposed algorithm works rather well even in the case of very close and narrow distributions $p(K)$ that creates an encouraging perspective for this method in investigating the complicated real systems. It should be noted, that algorithm CAS gives the unacceptably broad and, then, unsatisfactory solution even in a case of handling the exact $f(pM)$ dependence (Fig. 9).

Table 1. Application of the maxent algorithm to data subjected to the Gaussian errors

$s_r = 0.01$[*]				
J	Δ_{min}	U	S_{max}	Main peaks / corresponding $p(\log K_j)$
10	$1.2 \cdot 10^{-2}$	$1.6 \cdot 10^{-3}$	1.4	
50	$9.2 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$	1.6	3.1/0.36; 5.4/0.36; 6.4/0.10
100	$9.1 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$	1.8	
$s_r = 0.02$				
J	Δ_{min}	U	S_{max}	Main peaks / corresponding $p(\log K_j)$
10	$2.5 \cdot 10^{-2}$	$3.3 \cdot 10^{-3}$	1.4	
50	$1.9 \cdot 10^{-2}$	$2.9 \cdot 10^{-3}$	1.3	3.2/0.19; 3.4/0.19; 5.6/0.51
100	$1.9 \cdot 10^{-2}$	$2.8 \cdot 10^{-3}$	1.9	
$s_r = 0.05$				
J	Δ_{min}	U	S_{max}	Main peaks / corresponding $p(\log K_j)$
10	$6.3 \cdot 10^{-2}$	$8.6 \cdot 10^{-3}$	1.4	
50	$5.2 \cdot 10^{-2}$	$7.2 \cdot 10^{-3}$	1.8	3.6/0.27; 3.8/0.20; 6.0/0.20
100	$5.2 \cdot 10^{-2}$	$7.1 \cdot 10^{-3}$	2.2	
$s_r = 0.10$				
J	Δ_{min}	U	S_{max}	Main peaks / corresponding $p(\log K_j)$
10	0.13	$1.7 \cdot 10^{-2}$	1.2	
50	0.11	$1.8 \cdot 10^{-2}$	1.6	4.4/0.19; 4.6/0.54
100	0.11	$1.8 \cdot 10^{-2}$	1.4	

[*] s_r are the relative standard deviations of errors introduced into the exact $f([M])$ values.

Table 2. Application of the maxent algorithm to data subjected to the random relative errors uniformly distributed within range $[-a; a]^*$

$a = 0.005$				
J	$\Delta_{\min}$	U	$S_{\max}$	Main peaks / corresponding $p(\log K_j)$
10	$2.9\cdot10^{-3}$	$4.5\cdot10^{-4}$	1.2	
50	$2.9\cdot10^{-3}$	$4.2\cdot10^{-4}$	1.6	3.0/0.31; 5.2/0.26; 6.2/0.16
100	$2.9\cdot10^{-3}$	$4.2\cdot10^{-4}$	1.8	
$a = 0.01$				
J	$\Delta_{\min}$	U	$S_{\max}$	Main peaks / corresponding $p(\log K_j)$
10	$5.8\cdot10^{-3}$	$8.9\cdot10^{-4}$	1.2	
50	$5.7\cdot10^{-3}$	$8.5\cdot10^{-4}$	1.6	3.0/0.28; 5.2/0.35; 6.2/0.10
100	$5.7\cdot10^{-3}$	$8.5\cdot10^{-4}$	2.0	
$a = 0.025$				
J	$\Delta_{\min}$	U	$S_{\max}$	Main peaks / corresponding $p(\log K_j)$
10	$1.4\cdot10^{-2}$	$2.2\cdot10^{-3}$	1.3	
50	$1.4\cdot10^{-2}$	$2.0\cdot10^{-3}$	1.5	3.0/0.22; 5.4/0.41; 5.6/0.14
100	$1.4\cdot10^{-2}$	$2.0\cdot10^{-3}$	1.6	
$a = 0.05$				
J	$\Delta_{\min}$	U	$S_{\max}$	Main peaks / corresponding $p(\log K_j)$
10	$2.9\cdot10^{-2}$	$4.4\cdot10^{-3}$	1.5	
50	$2.9\cdot10^{-2}$	$3.6\cdot10^{-3}$	1.5	3.2/0.27; 5.6/0.30; 5.8/0.22
100	$2.9\cdot10^{-2}$	$3.7\cdot10^{-3}$	1.6	

* the relative standard deviations $s_r = \dfrac{a}{\sqrt{3}}$.
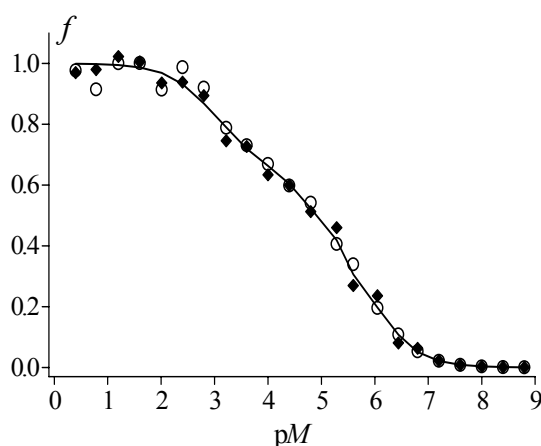
Fig. 4. Simulated binding degrees, f, versus pM. The solid line corresponds to the exact values, symbols o denote data with introduced Gaussian errors (relative standard deviation $s_r = 0.01$), symbols ♦ denote data with relative errors uniformly distributed within range [-0.025; +0.025].
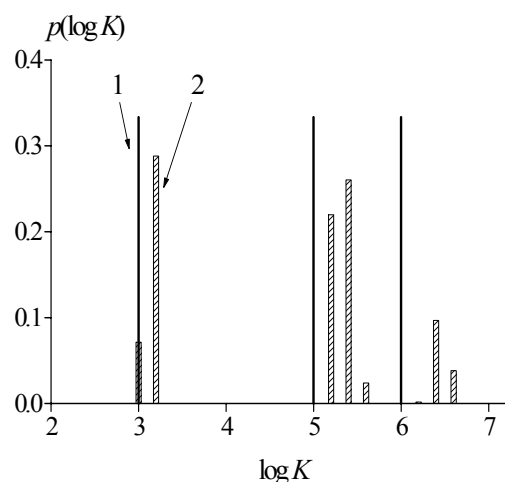
Fig. 5. The density function $p(K)$ calculated on the uniform grid with 50 knots with the aid of the maxent algorithm. 1 – the model density function; 2 – the function reconstructed from data with introduced Gaussian errors ($s_r = 0.01$).
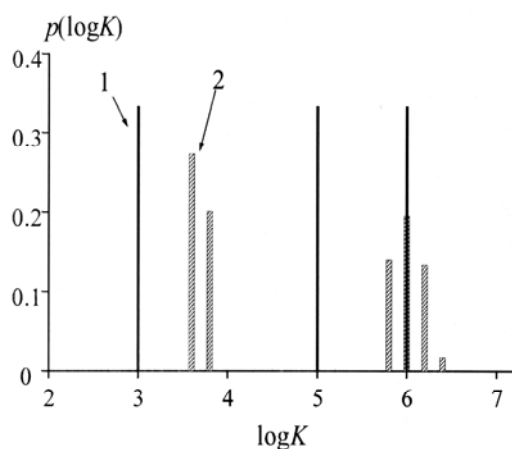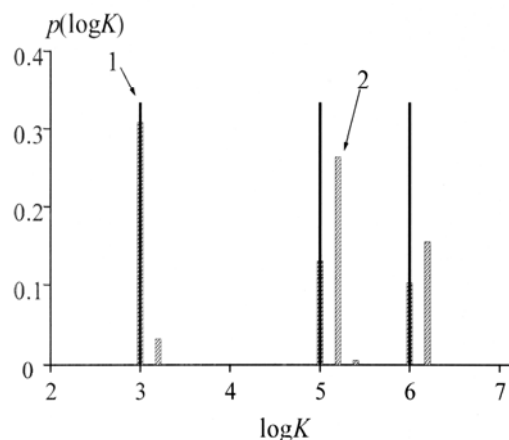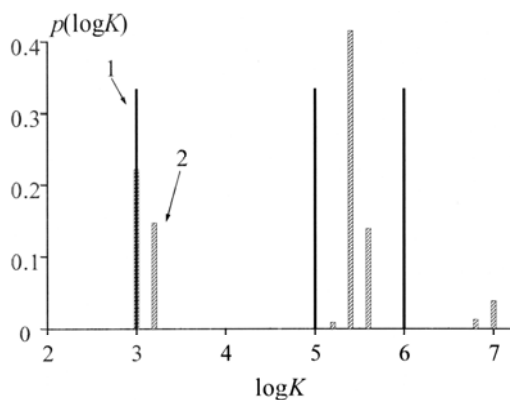
Fig. 6. The density function $p(K)$ calculated on the uniform grid with 50 knots with the aid of the maxent algorithm. 1 – the model density function; 2 – the function reconstructed from data with introduced Gaussian errors ($s_r = 0.05$).

Fig. 7. The density function $p(K)$ calculated on the uniform grid with 50 knots with the aid of the maxent algorithm. 1 – the model density function; 2 – the function reconstructed from noisy data (relative random errors were distributed uniformly within range [-0.005; +0.005]).

Fig. 8. The density function $p(K)$ calculated on the uniform grid with 50 knots with the aid of the maxent algorithm. 1 – the model density function; 2 – the function reconstructed from noisy data (relative random errors were uniformly distributed inside interval [-0.025; +0.025]).
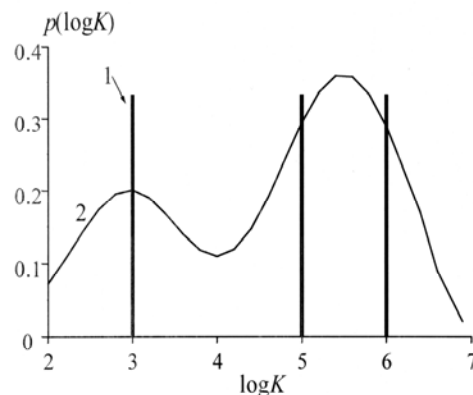
Fig. 9. The density function $p(K)$ calculated by algorithm CAS from the exact data on the uniform grid with 22 knots. 1 – the model density function; 2 – the reconstructed one.

Conclusions

1. The solution of the first kind integral equations is the ill-posed problem, and this results in definite computational difficulties. To overcome them, the most prospective algebraic numerical methods use (i) additional information which narrows a possible set of solutions to a compact; (ii) the Tikhonov α-regularization technique and (iii) the cross-validation technique. The proposed algorithm CAS applies to all these features. It is numerically stable in computing density and integral distribution functions; unlike its analogs, it is non-iterative and rapid.

2.   The method based on the maxent approach and developed as an alternative to the traditional procedures was shown to generate affinity distributions close to the model ones even in the case of narrow density functions. This feature of the method makes it prospective for studying complicated real systems.

References

1. Jaroniec M., Madey R. Physical Adsorption on Heterogeneous Solids. Amsterdam, e.a.: Elsevier, 1988.
2. Мамлеев В.Ш., Золотарев П.П., Гладышев П.П. Неоднородность сорбентов: (феноменологические модели). Алма-Ата: Наука, 1989.
3. De Wit J.C.M., Van Riemsdijk W.H., Koopal L.K. Environ. Sci. and Technol. 1993. V. 27. No 2. P. 2015-2022.
4. Taylor H.S. J. Phys. Chem. 1926. V. 30. No 21. P. 145-171.
5. Холин Ю.В. Количественный физико-химический анализ комплексообразования в растворах и на поверхности химически модифицированных кремнеземов: содержательные модели, математические методы и их приложения. Харьков: Фолио, 2000.
6. Yates D.E., Levine S., Healy T.W. J. Chem. Soc. Faraday Trans. Part I. 1974. V. 70. No 10. P. 1807-818.
7. Davis J.A., Leckie J.O. J. Colloid Interface Sci., 1980, V. 74. No 1. P. 32-43.
8. Джейкок М., Парфит Д.А. Химия поверхности раздела фаз: Пер. с англ. М.: Мир, 1984.
9. Zhmud B.V., Pechenyi A.B., Golub A.A. Functional Materials. 1995. V. 2. No 1. P. 44-50.
10. Charmas R., Piasecki W. Langmuir. 1996. V. 12. No 22. P. 5458-5465.
11. Hiemstra T., Van Riemsdijk W.H., Bolt G.H. J. Colloid Interface Sci. 1996. V. 183. No 1. P. 91-104.
12. Hiemstra T., Van Riemsdijk W.H. J. Colloid Interface Sci. 1996. V. 179. No 2. P. 488-508.
13. Koopal L.K. In: Adsorption on New and Modified Inorganic Sorbents / Ed. by A. Dabrowsky, V.A. Tertykh. Amsterdam e.a.: Elsevier, 1996. – P. 757-796.
14. Zhmud B.V. J. Colloid Interface Sci. 1996. V. 183, No 1. P. 111-117.
15. Budarin V.L., Pecheny A.B., Zaitsev V.N. Functional Materials. 1997. V. 4. No 4. P. 381-384.
16. Contescu C., Jagiello J., Schwarz J.A. Langmuir. 1993. V. 9. No 7. P. 1754-1765.
17. Темкин М.И., Левич М. Журн. физ. химии. 1946. Т. 20. No 12. C. 1441-1457.
18. Sips R. J. Chem. Phys. 1948. V. 16. No 5. P. 490-495.
19. Тодес О.М., Бондарева А.К., Хачкурузов Г.А. В Кн.: Проблемы кинетики и катализа. VII. Статистические явления в гетерогенных системах. М.-Л.: Изд-во АН СССР. 1949. C. 263-296.
20. Zeldowitch J. Acta Physicohimica U.R.S.S. 1935. V. 1. No 6. P.961-974.
21. Рогинский С.З. Адсорбция и катализ на неоднородных поверхностях. М.-Л.: Изд-во АН СССР, 1948.
22. Nederlof M., Van Riemsdijk W.H., Koopal L.K. J. Colloid Interface Sci. 1990. V. 135. No 2. P. 410-425.
23. Nederlof M., Van Riemsdijk W.H., Koopal L.K. Environ. Sci. Technol. 1994. V. 28. No 6. P. 1037-1044.
24. Högfeldt E. Arkiv för Kemi. 1962. V. 20. No 19. P. 215-224.
25. Jaroniec M. Surf. Sci. 1975. V. 50. No 2. P. 553-564.

26. Thakur A.M., Munson P.J., Hunston D.L., Rodbard D. Anal. Biochem. 1980. V. 103. No 2. P.240-254.
27. Яронец М., Осцик Я., Бройер П. В кн.: Адсорбенты, их получение, свойства и применение. Л.: Наука, 1985. С. 80-85.
28. Quiñones I., Guiochon G. J. Colloid Interface Sci. 1996. V. 183. P. 57-67.
29. Nederlof M. Analysis of binding heterogeneity. Ph. D. Thesis. Wageningen Agricultural University, Wageningen, 1992.
30. Roles J., Guiochon G. J. Phys. Chem. 1991. V. 95. No 10. P. 4098-4109.
31. Ross S., Morrison I.D. Surface Sci. 1975. V. 52. No 1. P. 103-119.
32. Sacher R.S., Morrison I.D. J. Colloid Interface Sci. 1977. V. 70. No 1. P.153-166.
33. Лоусон Ч., Хенсон Р. Численное решение задач метода наименьших квадратов: Пер. с англ. М.: Наука, 1980.
34. Phillips D.L. J. Assoc. Comput. Mach. 1962. V. 9. P. 84-97.
35. Baker C.T.H., Fox L., Mayers D.F., Wright K. Comput. J. 1964. V. 7. P. 141-148.
36. Hanson R.J. SIAM J. Numer. Anal. 1971. V. 8. No 3. P. 616-622.
37. Leunberger B., Schindler P.W. Anal. Chem. 1986. V. 58. No 7. P. 1471-1474.
38. Братская С.Ю., Голиков А.П. Журн. аналит. химии. 1998. Т. 53. No 3. С. 265-271.
39. Гармаш А.В., Устимова И.В., Кудрявцев А.В. и др. Журн. аналит. химии. 1998. Т. 53. No 3. С. 241-248.
40. Stanley B.J., Guiochon G. J. Phys. Chem. 1993. V. 97. No 30. P. 8098-8104.
41. Stanley B.J., Bialkowski S.E., Marshall D.B. Anal. Chem. 1993. V. 65. No 3. P. 259-267.
42. Karush F., Sonenberg M. J. Am. Chem. Soc. 1949. V. 71. No 4. P.1369-1376.
43. Perdue M., Lyttle C.R. Environ. Sci. and Techol. 1984. V. 17. No 11. P. 654-660.
44. Тихонов А.Н. Докл. АН СССР. 1963. Т. 153. No 1. С. 49-52.
45. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1986.
46. House W.A., Jaycock M.J. J. Colloid and Polym. Sci. 1978. V. 256. No 1. P. 52-61.
47. Верлань А.Ф., Сизиков В.С. Интегральные уравнения: методы, алгоритмы, программы. Киев: Наукова думка, 1986.
48. Stone M. J. of the Royal Stat. Soc. 1974. V. 36. No 2. P. 111-147.
49. Osten D.W. J. of Chemometrics. 1988. V. 2. P. 39-48.
50. Stanley B.J., Bialkowski S.E., Marshall D.B. Anal. Chem. 1993. V. 65. No 3. P. 259-267.
51. Craven P., Wahba. G. Numer. Mathematics. 1979. V. 31. P. 377-403.
52. Wahba G. In: CBMS-NSF Regional Conf. Series in Appl. Math. Chapter 4. Philadelphia: Soc. for Industrial and Appl. Mathematics, 1990. P. 45-65.
53. Титаренко В.Н., Ягола А.Г. Вестник Московского ун-та. 2000. Сер. 3. Физика, Астрономия. No 4. С. 15-18.
54. Леонов А.С., Ягола А.Г. Вестник Московского ун-та. 1995. Сер. 3. Физика, Астрономия. No 4. С. 28-33.
55. Адамсон А.А. Физическая химия поверхностей: Пер. с англ. М.: Мир, 1979.
56. Холин Ю.В., Мерный С.А. Журн. физ. химии. 1996. Т. 70. No 6. С.1065-1071.
57. Kholin Yu., Myerniy S., Varshal G. Ads. Sci. 2000. V. 18. No 3. P. 267-293.
58. Beraa A.K., Biliasb Y. J. of Econometrics. 2002. V. 107. P. 51-86.
59. Jaynes E.T. Phys. Rev. 1957. V. 106. P. 620–630.
60. Livesey A.K., Skilling J. Acta Cryst. 1985. V. A41. P. 113-122.
61. Wu X. J. of Econometrics. 2003. V. 115. P. 347-354.
62. Mead L.R. J. Math. Phys. 1986. V. 17. P. 2903-2907.
63. Amato U., Hughes W. Inverse Problems. 1991. V. 7. P. 793-808.

Приведен обзор численных методов расчета функций распределения сорбционных центров по константам сродства с компонентами растворов в модели биографической неоднородности сорбента. Проанализирована математически некорректная природа задачи и способы привлечения информации, позволяющей ограничить множество допустимых решений. Описаны разработанные авторами алгоритм CAS, основанный на решении условно корректной задачи расчета интегральной функции распределения, и алгоритм, модифицирующий применительно к рассматриваемой задаче метод максимума энтропии Джейниса.