

NEURAL NETWORK-BASED IDENTIFICATION OF NONLINEAR ADSORPTION ISOTHERMS

Weihoa Gao * Sebastian Engell *

** Process Control Laboratory, Department of Biochemical and
Chemical Engineering,
University of Dortmund, Emil-Figge-Str. 70, D-44221
Dortmund, Germany*

Abstract: The adsorption isotherm is the most important parameter in chromatographic separation process. In this paper, a numerical isotherm estimation method based on neural networks is proposed. As not the absolute values but the slopes of the isotherms are most important in process simulation and optimization, the slopes of the isotherm instead of the absolute values are taken as the outputs of the neural network. A tailored method is used to initialize the neural network. The parameters of the neural network are adjusted to minimize the difference between the simulated and the measured profiles. The issue of the design of the experiments, which decides the amount of the isotherm information contained in the profiles, is discussed. Simulation and experimental studies illustrate the potential of the proposed method.

Keywords: Chromatography, neural networks, nonlinear identification, adsorption isotherm.

1. INTRODUCTION

The adsorption isotherm describes the distribution of the solute between the two phases of a chromatographic system. Different components can be separated because they have different adsorption isotherms. In analytical applications of chromatography the concentration of the analyte in the mobile phase usually is very small. Therefore the assumption of linear isotherms is justified for these processes. In most preparative applications, however, the concentration range of interest is wide and therefore the linearity assumption is no longer valid. The functional relationship between the equilibrium concentrations in the two phases can become very complex, depending on the concentration range, the retention mechanism, and the phase system. Furthermore, in the case of multi-component mixtures, an additional complexity results from the competition between the different components for interaction with the stationary phase.

The identification of the adsorption isotherm is very important for applying model-based methods to the optimization and the operation of chromatographic separation processes. Adsorption isotherms can be determined by static and dynamic experimental methods (Ruthven, 1984; Guiochon *et al.*, 1994). Static methods consist of near-equilibrium measurements of the extent of adsorption at specified, constant adsorbate concentrations. The extent of adsorption has been measured by methods including microbalance gravimetry, infrared absorption, and thermogravimetry. Static methods are tedious and time-consuming and require a large amount of material. Dynamic methods consist of frontal analysis (FA) (James and Phillips, 1954; Schay and Szekeley, 1954; Glückauf, 1955), frontal analysis by characteristic points (FACP) (Glückauf, 1947), peak-maxima methods (Kipping and Winter, 1965), elution by characteristic point (ECP) (Cremer and Huber, 1961), and step-and-pulse (minor disturbance) methods (Helfferich and Peterson, 1963). Among them, the frontal analy-

sis (FA) and step-and-pulse can be extended to measure multicomponent isotherms. The simple wave method can also be used for the determination of binary isotherms in cases in which the deviation of isotherm from the Langmuir model is moderate (Ma *et al.*, 1990). These experimental methods need large volumes of concentrated solutions. This is expensive for valuable substances. Moreover, the long experimental period is another drawback. An alternative approach is to approximate the adsorption isotherm numerically. The isotherm functional form is chosen at the outset. Then by minimizing the difference between calculated and experimental profiles, the parameter values are adjusted until a best fit to the experimental data is found. Dose *et al.* determined the equilibrium isotherms of N-benzoyl-(D,L)-alanine and N-benzoyl-(D,L)-phenylalanine on immobilized bovine serum albumin from the peak shapes of single component (Dose *et al.*, 1991). They used a modified simplex algorithm to find the best parameters of Bi-Langmuir isotherms and found a good agreement with the isotherms determined with frontal analysis. James and Sepúlveda developed a more sophisticated algorithm for the estimation of the isotherm parameters by the inverse method (James and Sepúlveda, 1994; James *et al.*, 1999). They used the conjugate gradient algorithm for the minimization of the objective function that - besides the conventional least squares - took into account the difference between the first moments of the peaks, as well. This algorithm was applied to the estimation of the competitive Bi-Langmuir and Moreau isotherms of Ketoprofen enantiomers on a Chiracel OJ column, as well as of the competitive Langmuir isotherm of benzyl alcohol and 2-phenylethanol on a C_{18} column from individual band profiles. The competitive isotherms estimated with the inverse method agreed well with the data obtained from conventional methods. Felinger *et al.* derived the parameters of competitive isotherms from the band profiles of a racemic mixture of the 1-phenyl-1-propanol enantiomers by means of a nonlinear least-squares method (Felingner *et al.*, 2003). The isotherms determined from the overloaded band profiles matched extremely well the isotherms determined by frontal analysis.

The numerical methods transfer the experimental effort to the computational effort. Compared with the experimental methods, they are more economical because only a small amount of concentrated solution is used. At the same time, the computational time is shortened by using high-speed computers. However

the functional form of the isotherm must be assumed prior to the numerical approximation. If there is a structural mismatch between the assumed functional form and the actual isotherm, the numerical methods cannot guarantee a good result. In practice, several isotherm forms should be tried in order to estimate the most suitable model. On the other hand, a universal form of the isotherm should be used when there is no knowledge about the isotherm. This leads to the idea of using a neural network to represent the isotherm model and of estimating the neural network parameters from the experimental band profiles. This idea was first described in (Daniel *et al.*, 2002). In this paper, a systematic way to use neural network to identify isotherm is introduced.

2. THEORY

2.1 Chromatographic separation process

In a batch chromatographic separation process, one charge of the mixture to be separated is injected into the column together with a suitable solvent. The charge is carried through the column by continuously injecting additional desorbent. Due to the different adsorption affinities, the different components have different migration velocities and are gradually separated while moving along the column. The concentrations of the components are measured at the outlet of the column (Fig. 1).

The mathematical modelling of the chromatographic separation process has been extensively described in the literature by several authors, and is in most cases based on a differential mass balance. The most general 1-dimensional model, which takes into consideration all important effects, is the general rate model:

$$\begin{aligned} \frac{\partial c_i}{\partial t} + u \frac{\partial c_i}{\partial x} + \frac{3(1-\varepsilon)k_{l,i}}{\varepsilon r_p} (c_i - c_{p,i}(r_p)) &= D_{ax,i} \frac{\partial^2 c_i}{\partial x^2} \\ (1-\varepsilon_p) \frac{\partial q_i}{\partial t} + \varepsilon_p \frac{\partial c_{p,i}}{\partial t} &= \varepsilon_p D_{p,i} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial c_{p,i}}{\partial r} \right) \right] \end{aligned} \quad (1)$$

These two partial differential equations describe the concentrations in the mobile phase (c_i) and in the stationary phase (q_i and $c_{p,i}$). Adsorption isotherms are used to relate the concentrations q_i (in the solid of the stationary phase) and $c_{p,i}$ (in the liquid of the stationary phase):

$$q_i = f_i(c_{p,1}, c_{p,2}, \dots, c_{p,i}, \dots, c_{p,n}). \quad (2)$$

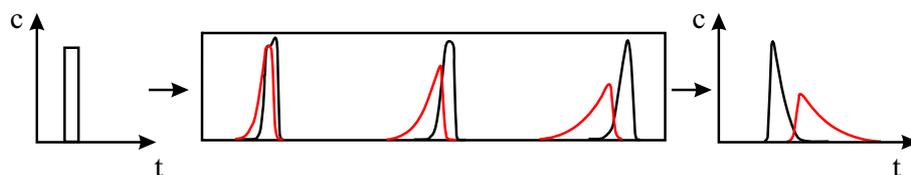


Fig. 1. Illustration of a chromatographic separation process.

It should be noted that in the isotherm equations, the concentration $c_{p,i}$ is usually denoted by c_i . This does not mean that they are the same. An efficient numerical solution for the general rate model incorporating arbitrary nonlinear isotherms was proposed by Gu (Gu, 1995). The mobile phase and the stationary phase are first discretized using the finite element and the orthogonal collocation methods respectively. The resulting ODE system is solved using an ODE solver that is based on the Gear's method. The numerical solution yields the concentrations of the components in the column at different locations and times. The concentrations at the outlet of the column are used to generate the band profiles. From the model (1), it is obvious that the derivatives $\partial q_i/\partial c_j$ and not the absolute values of the isotherm are critical.

2.2 Neural network isotherm

Two-layer neural networks with sigmoidal neurons in the hidden layer and linear neurons with bias in the output layer are used to describe the isotherms. The outputs of the neural networks are the slopes of the isotherms $\partial q_i/\partial c_j$, instead of the values q_i . Fig. 2 depicts the structure of a simple neural network for binary chromatography.

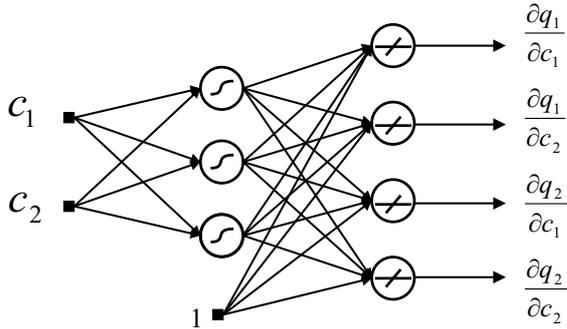


Fig. 2. Structure of neural network isotherm model.

2.3 Identification of the neural network isotherm

The identification principle is shown in Fig. 3. The neural network isotherm is used in the general rate model to simulate the band profiles. The experimental band profiles are compared with the simulated band profiles by evaluating the distance function

$$\sum_i \sum_j (c_{i,j}^{sim} - c_{i,j}^{meas})^2, \quad (3)$$

where i is the index of component, j is the index of the sampling points. The parameters of the neural network isotherm are adjusted to minimize the distance, using the Levenberg-Marquardt method with line-search.

A simple method is proposed to initialize the neural network isotherm. The weights of the input layer are set to random numbers in $(0,1)$. The weights of the

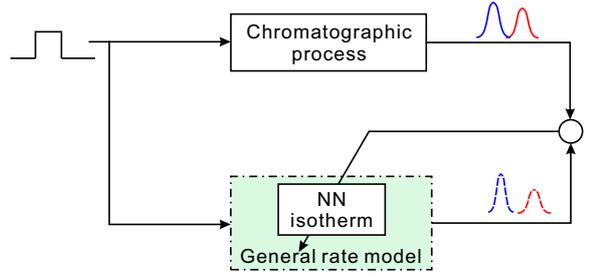


Fig. 3. Principle of neural network-based isotherm identification.

output layer are set to 0. The biases of $\partial q_i/\partial c_j$ ($i = j$) are set by a trial and error procedure. They are adjusted until the simulated band profiles have roughly the same position as the experimental band profiles. The biases of $\partial q_i/\partial c_j$ ($i \neq j$) are set to arbitrary negative numbers in $(-1,0)$. The steps in the solution of the identification problem are as follows:

- (1) Analysis of the chromatographic process and design of proper experimental conditions, i.e. the operating parameters, the input concentration, and the choice of multi injection or single injection. Running the experiment and measuring the band profiles.
- (2) Selection of the neural network structure. Besides the above structure, other more complex structures are also possible.
- (3) Choice of the number of neurons in the hidden layer. The number of the neurons in the hidden layer decides the ability of the neural network to approximate a nonlinear function. An iterative procedure is used to determine this number.
- (4) Estimation of the parameters using the Levenberg-Marquardt method with line-search.
- (5) Validation of the neural network isotherm by comparing simulations of eq. (1) to measured data.

3. RESULTS AND DISCUSSION

The neural network isotherm was first checked in a simulation study, and then it was applied to estimate isotherms from experimental data.

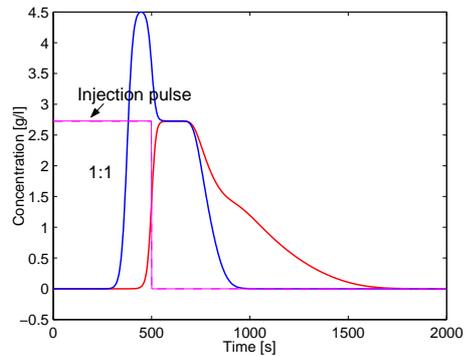


Fig. 4. Overloaded band profiles of the single injection, 1:1 feed mixture.

3.1 Simulation study

In the simulation study, an separation process which exhibits a strongly nonlinear adsorption isotherm was considered (Hanisch, 2002). The competitive adsorption to the solid phase of the two components is described by a Bi-Langmuir isotherm:

$$q_1 = \frac{0.0127c_1}{1 + 1000c_1 + 1008c_2} + \frac{23.412c_1}{1 + 333.1c_1 + 94.5c_2}$$

$$q_2 = \frac{8.979c_2}{1 + 211.4c_1 + 43.46c_2} \quad (4)$$

By simulating the general rate model with the Bi-Langmuir isotherm band profiles were acquired. They are regarded as the "experimental" band profiles in the identification procedure. Fig. 4 shows the band profiles for an overloaded injection. Different numbers of neurons in the hidden layer were tried. The results are listed in table 1. From this, the number of the neurons in the hidden layer is set to three.

Table 1. Estimation results for different numbers of neurons in the hidden layer

Number of neurons in the hidden layer	Residual $\times 500$	Iterations	Function count
1	9.2085	29	438
2	0.1886	33	714
3	0.1184	40	1087
4	0.0409	49	1625

The estimated neural network isotherms were compared with the original Bi-Langmuir isotherm in Fig. 5. The neural network isotherm approaches the

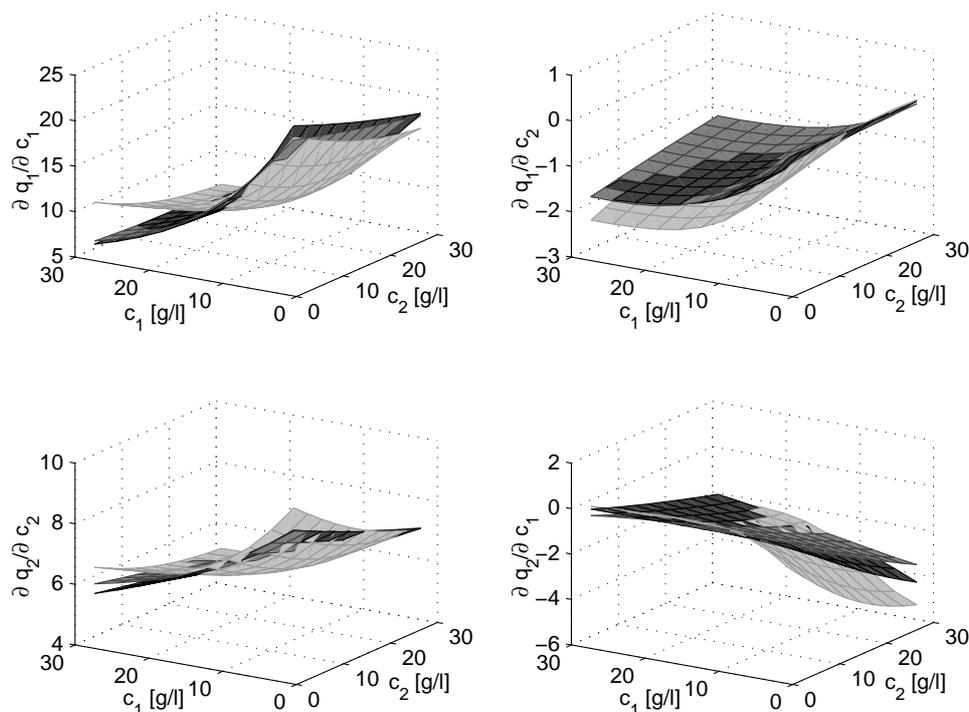


Fig. 5. Comparison of the Bi-Langmuir isotherm and the estimated neural network isotherms, black surface: Bi-Langmuir isotherm, light grey surface: neural network isotherm (single injection), dark grey surface: neural network isotherm (multi-injection).

Bi-Langmuir isotherm in all four derivatives. The difference between them can be explained as insufficient isotherm information contained in the "experimental" band profiles. To show this, all the combinations of concentrations which are involved in the numerical solution of the model (1) are depicted in Fig. 6. The "experimental" band profiles are only the concentrations at the outlet of the column. Most of the points are along a closed curve in the concentration space. To improve the neural network isotherm, the experiment should be specially designed so that enough information is contained in the band profiles. As a first attempt, a multi injection was simulated. The injection consisted of three pulses, as shown in Fig. 7. For every pulse, there is different feed concentration ratio (1:1, 3:1 and 1:3). The concentration combinations in the numerical solution are shown in Fig. 8. Compared

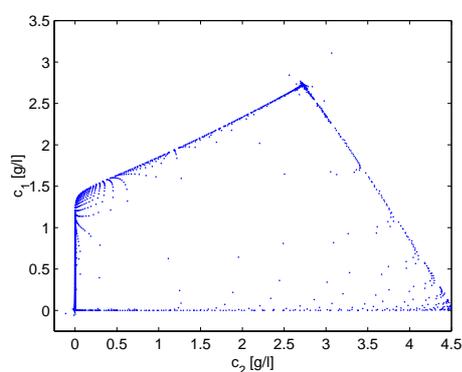


Fig. 6. Concentration combinations in the column, single injection, 1:1 feed mixture.

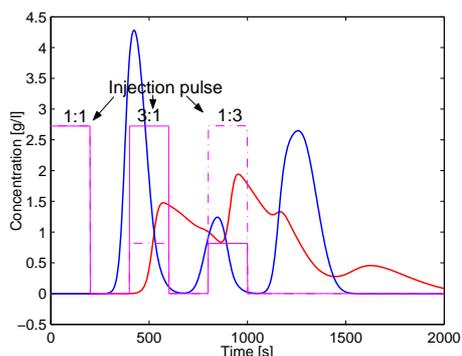


Fig. 7. Band profiles of the multi injection, (1:1), (3:1) and (1:3) feed mixture.

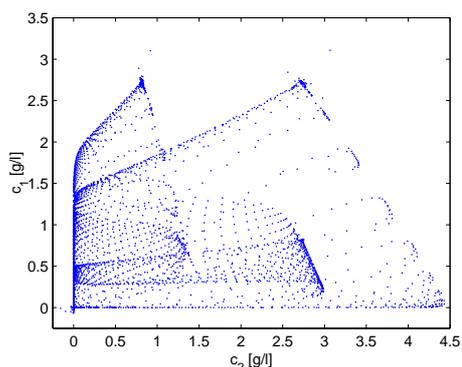


Fig. 8. Concentration combinations in the column, multi-injection.

with the single overloaded injection, a much larger area is covered. The same neural network isotherm model (3 neurons in the hidden layer) was used in the identification procedure. As Fig. 5 shows, the neural network isotherm model was remarkably improved.

3.2 Estimating from experimental data

The experimental data was supplied by the chair of process and plant design (Lehrstuhl für Anlagentechnik, Universität Dortmund, Germany). The separation of EMD53986 on a chiral stationary phase is considered. For the measurement of the EMD53986 a preparative UV detector cell was used. The band profiles were calculated from the UV signals.

The experimental data consists of five sets of band profiles for different injection volumes. The neural network isotherm was identified from the most overloaded band profiles (Fig. 9) and was tested using the remaining band profiles. The results are shown in Fig. 9 and Fig. 10, and compared with an optimized Bi-Langmuir isotherm. In Fig. 9, the simulated band profiles using the neural network isotherm match the experimental band profiles well, while with the Bi-Langmuir isotherm large errors are seen because of its fixed structure. In Fig. 10, the neural network isotherms give a better prediction in the overlapping

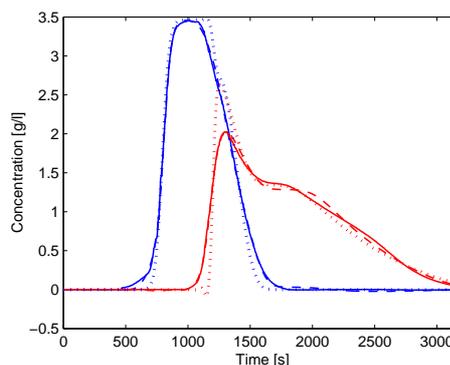


Fig. 9. Comparison of measurement and estimation for injection volume of 120 ml, solid lines: measurement, dashed lines: estimation of neural network isotherm (4 neurons in the hidden layer), dotted lines: estimation of Bi-Langmuir isotherm.

region than the Bi-Langmuir isotherm. This is important for model-based control and optimization because the purities of the products are determined by the overlap of the profiles.

4. CONCLUSIONS

Numerical isotherm identification methods have attracted considerable attention in preparative chromatography. This paper studies the possibility of using neural networks as the isotherm functional form in the numerical methods. The isotherm slopes are generated as the outputs of a neural network. By increasing the number of hidden layer neurons the neural network can represent any adsorption isotherm. The simulation and experimental studies prove that adsorption isotherms can be estimated well from band profiles by using neural networks as the isotherm functional form. The key issue is to design the experiment to generate enough information for the identification. The optimization of the experiment will be the focus of future research.

ACKNOWLEDGEMENTS

The work of the first author was supported by a scholarship of the NRW Graduate School of Production Engineering and Logistics at Universität Dortmund. This support is very gratefully acknowledged.

REFERENCES

- Cremer, E. and H. F. Huber (1961). Measurement of adsorption isotherms at high temperature with the help of gas-solid elution chromatography. *Angew. Chem.*, **73**, 461-465.
- Daniel, G., J. Dienstuhl, S. Engell, S. Felske, K. Goser, R. Klinkenberg, K. Morik, O. Ritthoff and H. Schmidt-Traub (2002) Novel learning tasks, optimization, and their application. In: *Advances*

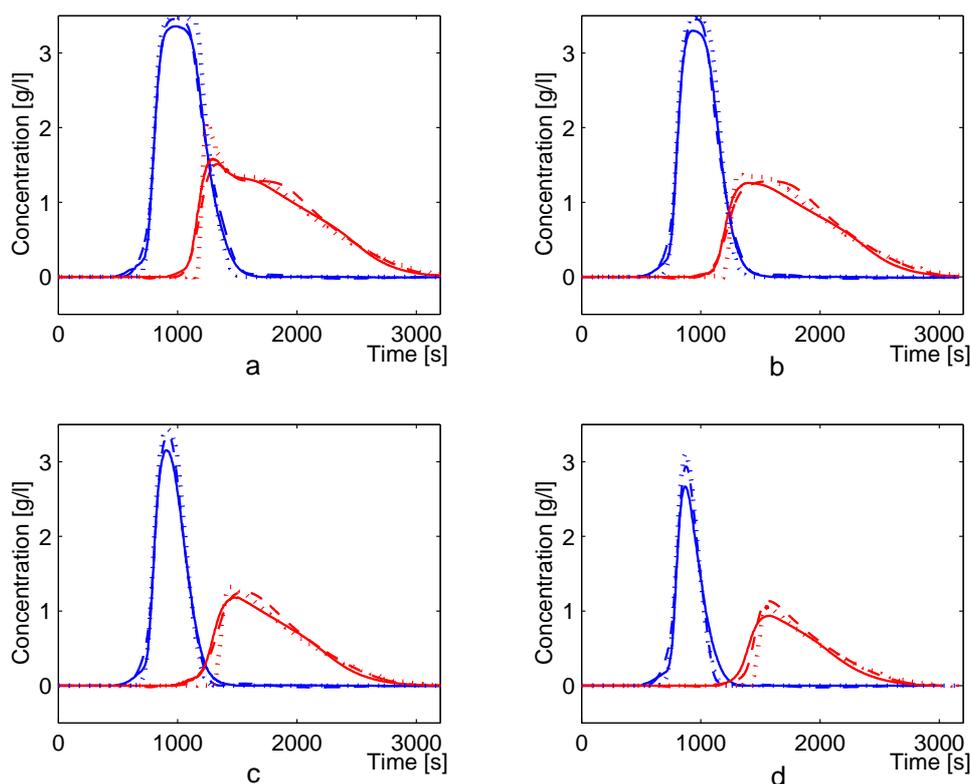


Fig. 10. Model validation for injection volumes of 100 ml, 80 ml, 60 ml and 40 ml, solid lines: measurements, dashed lines: simulation of neural network isotherm, dotted lines: simulation of a Bi-Langmuir isotherm.

- in Computational Intelligence: Theory and Practice* (H.-P. Schwefel, I. Wegener and K. Weinert, Ed.). Chap. 8, pp. 278-294. Springer Verlag, Berlin.
- Dose, E. V., S. Jacobson and G. Guiochon (1991). Determination of isotherms from chromatographic peak shapes. *Anal. Chem.*, **63**, 833-839.
- Felinger, A., A. Cavazzini and G. Guiochon (2003). Numerical determination of the competitive isotherm of enantiomers. *Journal of Chromatography A*, **986**, 207-225.
- Glückauf, E. (1947). Theory of chromatography: Part II. Chromatography of a single solute. *J. Chem. Soc.*, 1302-1308.
- Glückauf, E. (1955). Theory of chromatography: Formulas for diffusion into spheres and their application to chromatography. *Trans. Faraday Soc.*, **51**, 1540.
- Gu, T. (1995). *Mathematical Modelling and Scale Up of Liquid Chromatography*. Springer Verlag, New York.
- Guiochon, G., S. Golshan-Shirazi and A. Katti (1994). *Fundamentals of Preparative and Nonlinear Chromatography*. Academic Press, Boston.
- Hanisch, F. (2002). Prozessführung Präparativer Chromatographieverfahren (Operation of preparative chromatographic processes). Dr.-Ing. Dissertation. Shaker Verlag, Aachen. (in German)
- Helfferich, F. G. and D. L. Peterson (1963). Chromatographic methods for sorption isotherms and phase equilibria. *Science*, **142**, 661-662.
- James, D. H. and C. S. G. Phillips (1954). The chromatography of gases and vapours: Part III. the determination of adsorption isotherms. *J. Chem. Soc.*, 1066-1070.
- James, F. and M. Sepúlveda (1994). Parameter identification for a model of chromatographic column. *Inverse Problems*, **10**, 1299-1314.
- James, F., M. Sepúlveda, F. Scharton and I. Quinones (1999). Determination of binary competitive equilibrium isotherms from the individual chromatographic band profiles. *Chemical Engineering Science*, **54**, 1677-1696.
- Kipping, P. J. and D. G. Winter (1965). The measurement of adsorption isotherms by a gas chromatographic technique. *Nature*, **205**, 1002-1003.
- Ma, Z., A. Katti, B. Lin and G. Guiochon (1990). Simple wave effects in two-component nonlinear liquid chromatography. application to the measurement of competitive adsorption isotherms. *J. Phys. Chem.*, **94**, 6911-6922.
- Ruthven, D. M. (1984). *Principles of Adsorption and Adsorption Processes*. Wiley, New York.
- Schay, G. and G. Szekely (1954). Gas adsorption measurements in flow systems. *Acta Chim. Hung.*, **5**, 167-182.