

FEATURES FOR THE DETERMINATION OF PARAMETERS OF EQUILIBRIUM ADSORPTION MODEL BY A NONLINEAR REGRESSION PROCESS OF EXPERIMENTAL DATA

Ch. Chilev¹, I. Pentchev¹, F. Darcrim-Lamari², E. Simeonov¹

¹University of Chemical Technology and Metallurgy
8 Kl. Ohridski, 1756 Sofia, Bulgaria

²LSPM, CNRS, 99 Av. J.-B. Clément 93430
Villetaneuse, France
E-mail: cchilev@abv.bg

Received 11 July 2012
Accepted 12 October 2012

ABSTRACT

A mathematical model with four parameters describing the adsorption equilibrium of supercritical gases on microporous adsorbents in a wide range of pressures is studied. Numerical experiments related to determining the model parameters by nonlinear regression process of experimental data are carried out. From a physical point of view for a particular system adsorbent/adsorbate at a given temperature, these parameters are strictly defined. There was instability in determining the two of the parameters during the process of nonlinear regression related to proposed different initial values and obtained incorrect final results. The mathematical model is transformed by combining of these two parameters. Thus a three parametric model is received. The parameters of the resulting model are correctly determined during the process of nonlinear regression independently of the given initial values.

Keywords: adsorption, regression process, equilibrium.

INTRODUCTION

The adsorption processes have found widespread application in modern technologies. They are used for fine purification of industrial gases, air, and to obtain components with high purity. Particularly effective is the use of adsorption processes for taking harmful or valuable components with low concentrations from mixtures thereof during various technological processes [1]. Important point in the design of adsorption processes takes into account proper description and modeling of equilibrium of the studied systems, which sets as its base for further calculations. One fairly recent models describing adsorption equilibrium of supercritical gases on microporous adsorbents in a wide range of pressures is that proposed by Ch.Chilev et al.[2, 3]. The model

based on a combination between the potential theory firstly proposed by Polany and later developed by Dubinin, the concept of lattice structure of the gas phase and using mathematical formalism of statistical thermodynamics and quantum mechanics. As a result, a four-parametric equation of the following type is received [2]:

$$\Gamma = \Gamma_{\max} \left[\frac{bx}{1-x} e^{-ax} \right]^l \quad (1)$$

Eq. (1) relates the Gibbs mass excess to the variable x (degree of occupation of the lattice in the gas phase), respectively, the pressure in the gas phase [2]. The equation contains four parameters, namely:

- Maximum Gibbs mass excess - Γ_{\max} ,
- constant indicating the ratio of adsorbed phase volume V_s to the gas phase volume V_g - $b = V_s/V_g$

- constant energetic

$$a = \frac{\delta_g E_{m-m}}{2kTM_g}$$

(reporting the intermolecular interactions),

- constant energetic

$$t = \frac{kT}{E_{m-s}}$$

(reporting the interactions between the adsorbed molecules and the solid phase).

In the above expressions: E_{m-m} - energy of interaction between gas molecules relation to a single molecule, J ; δ_g - coordination number of the gas phase lattice; M_g - total number of cells in the gas phase lattice (the number of independent active centers); k - constant of Boltzmann, $m^2kg/(Ks^2)$; T - temperature, K .

The degree of occupation of the gas phase lattice is the ratio between the number of filled cells and the total number of cells in this phase, and can be determined by the expression [2]:

$$x = \frac{\rho}{\rho^{\max}} \frac{N_A}{M_m} \quad (2)$$

In the equation (2) N_A is the Avogadro number, M_m [g/mol] is the gas molar mass. The maximum density of the closed system corresponding to $x = 1$ is $\rho^{\max} = \sqrt{2}/\sigma^3$, where the radius of the gas molecules is σ . Using one of the proposed in the literature steady state equations for real gas (PVT models) the mass density of the gas ρ can be calculated. In the present work for determining the mass density of the gases the model of Peng Robitson is used [4].

NUMERICAL EXPERIMENTS

Using nonlinear regression of experimental data the model parameters are determined. To perform the nonlinear regression procedure both the program function “*datafit*” in software environment SciLab4.0 or sub-optimization program (curve fitting toolboxes) in the software environment MatLab9.0 [5-7] was used. The software environment MatLab9.0 has much greater opportunities for optimization and control during the nonlinear regression process.

The residuals from a fitted model are defined as the differences between the response data and the fit to the response data at each predicted value [8].

$$\text{residual} = \text{data} - \text{fit}$$

Mathematically, the residual for a specific predicted value is the difference between the response value y and the predicted response value \hat{y} .

$$r = y - \hat{y}$$

GOODNESS OF FIT STATISTICS

The Curve Fitting Toolbox supports the goodness of fit statistics for parametric models:

- The sum of squares due to error (SSE)
- R-square
- Adjusted R-square
- Root mean squared error (RMSE)

Sum of Squares Due to Error. This statistics measures the total deviation of the response values from the fit to the response values. It is also called the summed square of residuals and is usually labeled as SSE [9].

$$SSE = \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2$$

A value closer to 0 indicates a better fit. Note that the SSE was previously defined in the Least Squares Fitting Method.

R-Square. This statistics measures how successful the fit is in explaining the variation of the data [8, 9]. Put in a different way, R-square is the square of the correlation between the response values and the predicted response values. It is also called the square of the multiple correlation coefficient and the coefficient of multiple determination. R-square is defined as the ratio of the sum of squares of the regression (SSR) and the total sum of squares (SST). Here SSR is defined by the equation:

$$SSR = \sum_{i=1}^n w_i (\hat{y}_i - \bar{y})^2$$

SST is also called the sum of squares about the mean, and is defined as:

$$SST = \sum_{i=1}^n w_i (y_i - \bar{y})^2$$

where $SST = SSR + SSE$. Given these definitions, R-square is expressed as:

$$R - \text{square} = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

R-square can take on any value between 0 and 1, with a value closer to 1 indicating a better fit. For example, an R² value of 0.8234 means that the fit explains 82.34 % of the total variation in the data about the average.

If you increase the number of fitted coefficients in the model, R-square might increase although the fit may not improve. To avoid this situation, you should use the degrees of freedom adjusted R-square statistics.

Note that it is possible to get a negative R-square for equations that do not contain a constant term. If R-square is defined as the proportion of variance explained by the fit, and if the fit is actually worse than just fitting a horizontal line, then R-square is negative. In this case, R-square can not be interpreted as the square of a correlation.

Degrees of Freedom Adjusted R-Square. This statistics uses the R-square statistic defined above, and adjusts it based on the residual degrees of freedom. The residual degrees of freedom is defined as the number of response values n minus the number of fitted coefficients m estimated from the response values [9].

$$v = n - m$$

Here v indicates the number of independent pieces of information involving the n data points that are required to calculate the sum of squares. Note that if parameters are bounded and one or more of the estimates are at their bounds, then those estimates are regarded as fixed. The degrees of freedom are increased by the number of such parameters. The adjusted R-square statistics is generally the best indicator of the fit quality when you add additional coefficients to your model.

$$\text{adjusted } R - \text{square} = 1 - \frac{SSE(n-1)}{SST(v)}$$

The adjusted R-square statistics can take on any value less than or equal to 1, with a value closer to 1 indicating a better fit. Negative values can occur when the model contains terms that do not help to predict the response.

Root Mean Squared Error. This statistics is also known as the fit standard error and the standard error of the regression [9]. It is an estimate of standard deviation of the random component in the data.

$$RMSE = s = \sqrt{MSE}$$

where MSE is the mean square error or the residual mean square.

$$MSE = \frac{SSE}{v}$$

A RMSE value closer to 0 indicates a better fit.

RESULTS AND DISCUSSION

From mathematical point of view the process of nonlinear regression is very sensitive to the initial values of the parameters [10]. In this context, when we describe the experimental data using the model equation in the form of eq. (1), for two of the parameters one obtain very accurate results, i.e. regardless of the initial approximations values of these parameters, their final values are the same.

There is a problem with the other two parameters. Their final values depend much on the initial approximation values, and during the process of regression significantly changed. In Table 1 the initial and final values of the model parameters obtained during the process of nonlinear regression of equilibrium experimental data of ethylene on activated carbon [3] are given. The results show that for parameters a and t despite the different initial approximation values the same final results are obtained. Thus, these parameters are uniquely determined with sufficient precision.

However the final results for the other two parameters Γ_{\max} and b , very strongly depend on their initial approximation values. In setting the initial value 8 of Γ_{\max} , the obtained result is 11.466 and for initial value 10, respectively 4.98, i.e. the final values strongly depend on the initial approximations. For the parameter b the same result is obtained. Therefore, to be used correctly this model in the form of equation (1) is necessary to know the initial approximation values of both two parameters (Γ_{\max} and b) with great precision because during the regression process they are optimized somehow, but not calculated correctly.

Parameter $b = V_s/V_g$ is the ratio of the volume of adsorbate to the volume of gas phase per gram of the adsorbent. It is a constant for a given system adsorbents/adsorbate at fixed temperature. Since the proposed model is a new one and has never been used

Table 1. Parameters of the model.

Parameters	First regression		Second regression	
	initials	finals	initials	Finals
Γ_{\max}	8	11,465826	10	4,9849601
b	10	10,821713	100	99,465201
a	15	35,493338	35	35,61064
t	0.6	0,3754962	1,2	0,3980201

before, no data can be found for this parameter for different systems.

In terms of the maximum adsorbed amount (Γ_{\max}) for a given adsorbent/adsorbate system, in the literature the necessary data can be found. But there are differences in the results of Γ_{\max} given by different authors, depending on their use of adsorption models to describe the equilibrium. Thus, the existing results for this parameter are not really sure. Besides, the number of systems adsorbents/adsorbate for which can be found these results are not many.

Table 2 presents the results of different types of fit statistics produced during the nonlinear regression process of experimental data. The table shows that different initial values for the parameters, do not influence the resulting errors (the results from the first and second regression are the same). Thus, the initial approximations values of the parameters only influence the final results for two of the parameters, but not on the stability of the nonlinear regression process.

It is noteworthy that for both of the used initial approximation values of the parameters: the same product $K = \Gamma_{\max} b^t$ is obtained. If from Table 1 we calculate this product (K) at different initial values for the parameters, then in both cases the same numbers are obtained.

$$K = \Gamma_{\max} b^t = \begin{cases} 11,435826 \times 10,821713^{0,3755} = 28,0399 \\ 4,9849601 \times 99,465201^{0,3980} = 28,0399 \end{cases}$$

Therefore, we propose to combine the two parameters Γ_{\max} and b , thus eq. (1) is converted into the type

$$\Gamma = K \left[\frac{x}{1-x} e^{-ax} \right]^t \tag{3}$$

Here the new parameter is $K = \Gamma_{\max} b^t$. Thus, a three parameter model whose parameters are uniquely and certainly determined during the process of nonlinear regression is obtained. The values of the parameters obtained from new regression process for the same system adsorbent/adsorbate using eq. (3) are presented in Table 3.

The table shows that for the parameters a and t the same value as in Table 1 are obtained. Thus they are uniquely and certainly determined. For the third parameter K independent of initial approximation values, the final results are equal. Hence it can be concluded that the determination of this parameter using the model form of eq. (3) is correct.

In Table 4 the different types of fit statistics produced during the nonlinear regression of experimental data using eq. (3) are presented. The table shows that despite the initial approximations values of the model

Table 2. Fit statistics.

Statistic	First regression	Second regression
SSE	0.01743	0.01695
R-square	0.9728	0.9804
Adjusted R-square	0.8893	0.8324
RMSE	0,425627	0,433842

Table 3. Parameters of the model.

Parameters	First regression		Second regression	
	Initial	Final	Initial	Final
K	5	28,039962	100	28,039962
a	15	35,493225	5	35,493225
t	0.6	0,3754953	1,2	0,3754953

parameters, the same values for the regression errors (results of first and second regression are practically identical) are obtained. Thus, using eq. (3), the initial approximations for the parameters do not affect over the stability of regression analysis, which is a criterion for obtaining correct results.

Comparing Tables 2 and 4 (using the model equation in the form (1) and (3)) shows that the second case we have better fit statistics of regression. The difference in both cases for the errors: sum of squares due to error (SSE) and the square deviation (R-Square or R²), is very small. This indicates that the type of model equations ((1) or (3)) does not affect over these two statistical errors.

There is a change in the standard deviation (Root mean squared error RMSE). This statistics is fundamental and is related to the accuracy of regression, when its value is closer to zero indicates a better fit. In the first case (using the eq. (1), Table 2) RMSE H^o 0.43, while the second (using the eq. (3), Table 4) RMSE H^o 0.29. Thus, from view point of the regression process is more suitable use of the model equation in the form (3). It should be noted that if the number of fitted coefficients in the model increase, RMSE might increase although the fit may not improve. In our case, the numbers of parameters change. To avoid this situation, one should use the degrees of freedom adjusted R-square statistics (Adjusted R-Square). For this parameter value close to 1 is considered better. The comparison between Tables 2 and 4 shows that the values of this

parameter increase from 0.8893-0.8327 (in the first case) to 0.9974-0.9993 (in the second case). Thus, from view point of this parameter, in the second case the nonlinear regression process is better [12-13].

This shows that in terms of procedure for nonlinear regression of experimental data, it is better to use the model in the form (3). Thus it is not possible to determine the parameters Γ_{\max} and b , but only their proposed combination (parameter K). The resulting three parameters, K , a and t are uniquely and certainly obtained for a given system, and the results of regression analysis are largely reliable.

As a result from the above mentioned discussion, we recommend the following algorithm for treating of experimental data using eq. (3).

1. Gibbs excess adsorption is calculated by setting the values of temperature and pressure. With these values, the mass density of the bulk phase is deduced by the equation of state (PVP model) for the gas phase.

2. From the bulk mass density, one estimates the degree of occupation of the gas phase lattice (χ) using eq. 2.

3. Excess adsorption can be then deduced by a nonlinear fitting from eq. 3.

CONCLUSIONS

The four-parametric model describing the adsorption equilibrium of pure gases on microporous

Table 4. Fit statistics.

Statistic	First regression	Second regression
SSE	0.01102	0.01065
R-square	0.9904	0.9904
Adjusted R-square	0.9993	0.9974
RMSE	0,2848876	0,2848876

adsorbents is studied. The numerical experiments related to determining the model parameters by nonlinear regression of experimental data are made. The influence of the initial approximations values of parameters on their final values and the stability of the overall regression process is studied. Two energetic parameters of the model α and t are uniquely and certainly determined during the nonlinear regression process, regardless of their initial approximations value. For the other two parameters was found that their final results depend strongly on the initial approximation, i.e. they are not uniquely determined during the nonlinear regression process. Therefore, the mathematical association of these parameters in the new parameter K is done. The three-parametric model, whose parameters (K , α and t) are determined uniquely and certainly during the nonlinear regression process is obtained. Thus, the final values of the parameters and the fit statistics of the regression essentially do not depend on the initial approximations.

REFERENCES

1. S. Barto, S. Dacey, D.F. Quinn, *Fundamental of adsorption*, Engineering Foundation, New York, 1984.
2. Ch. Chilev, I. Pentchev, B. Weinberger, J.Ph. Passarello, L.F. Darkrim, A new approach to describe high-pressure adsorption isotherms in subcritical and supercritical conditions, *AIChE J.*, **55**, 2009, 1793-1802.
3. Ch. Chilev, L.F. Darkrim, E. Kirilova, I. Pentchev, Comparison of gas excess adsorption models and high pressure experimental validation, *Chemical Engineering Research and Design*, Available online 10 April 2012, <http://dx.doi.org/10.1016/j.cherd.2012.03.012>.
4. I.S. Sandler, *Chemical and Engineering Thermodynamics*, New York, John Wiley & Sons, Inc., 1989.
5. R. G. Duggleby, A nonlinear regression program for small computers, *Anal. Biochem.*, **110**, 1981, 9-18.
6. R. I. Jennrich, M. L. Ralston, Fitting nonlinear models to data, *Annu. Rev. Biophys. Bioeng.*, **8**, 1979, 195-238.
7. W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Minimization or maximization of functions. Numerical recipes: the art of scientific computing*, Cambridge: Cambridge University Press, 1986, 289-293.
8. N. R. Draper, H. Smith, *An introduction to nonlinear estimation. Applied regression analysis*, New York, Wiley, 1981, 458-517.
9. F. Mosteller, J. W. Tukey, *Data analysis and regression*, New York, Addison-Wesley, 1977.
10. H.J. Motulsky, L.A. Ransnas, Fitting curves to data using nonlinear regression: a practical and nonmathematical review, *The FASEB Journal*, **1**, 1987, 365-374.
11. M. S. Caceci, W. P. Cacheris, Fitting curves to data. The simplex algorithm is the answer, *Byte* 9, 1984, 340-362.
12. D. Ratkowsky, *Comparing parameter estimates from more than one data set. Nonlinear regression modelling: a unified and practical approach*, New York, Dekker, 1983, 135-152.