

PARAMETRIZATION OF ADSORPTION ISOTHERMS TO DATA  
COLLECTED BY ADSORBATE CONCENTRATION DIFFERENCE

By

S.R. Lindner,  
S.J. Schwager,  
and  
L.W. Lion

BU-1094-M

April 1990

**ESL-TR-90-05  
APRIL 1990**

**SORPTION EQUILIBRIA OF VAPOR PHASE ORGANIC POLLUTANTS  
ON UNSATURATED SOILS AND SOIL MINERALS**

**LEONARD W. LION, SAY KEE ONG, SHARON R. LINDNER,  
JEAN L. SWANGER, STEVEN J. SCHWAGER, AND  
TERESA B. CULVER**

**School of Civil and Environmental Engineering, Cornell University  
Ithaca, N.Y. 14853**

**PROJECT COMPLETION REPORT  
APPROVED FOR PUBLIC RELEASE  
DISTRIBUTION UNLIMITED**

**ENGINEERING AND SERVICES LABORATORY  
AIR FORCE ENGINEERING AND SERVICES CENTER  
TYNDALL AIR FORCE BASE, FLORIDA 32403**

## SECTION VI

### PARAMETRIZATION OF ADSORPTION ISOTHERMS TO DATA COLLECTED BY ADSORBATE CONCENTRATION DIFFERENCE

By: S.R. Lindner, S.J. Schwager and L.W. Lion

#### A. INTRODUCTION

Since the introduction of the nonlinear Langmuir adsorption isotherm in 1918, various curve-fitting techniques have been used to estimate the two parameters most frequently used to describe vapor-solid distribution, the monolayer capacity,  $\Gamma_m$ , and the adsorption affinity constant,  $c$ . Although graphically fitting a straight line by eye to a linearized form of this equation (as proposed by Langmuir) is still useful today, recent work by Kinniburgh and others has addressed regression techniques for parameterizing adsorption isotherms. Kinniburgh (1986) advocated using weighted nonlinear regression to overcome the difficulty of finding, or choosing, a linearized form of an isotherm equation, if such a linear form exists. Part of the argument for using weighted nonlinear least squares was that whenever possible, regression analysis should be performed directly on the data rather than on calculated quantities (as is necessary when an equation is linearized), to avoid uneven propagation of experimental error (Cabaniss and Shuman, 1988). Calculated quantities may also be introduced into regression analyses when the extent of adsorption is computed from concentration-level differences. This calculation is common in practice when extraction of adsorbate from the adsorbent phase is experimentally difficult.

This section examines the effectiveness and reliability of parameter estimation by nonlinear regression and compares it with an alternate numerical technique. In addition, isotherm equations have been rewritten in a form where the response variable,  $\Gamma$ , the amount of adsorption per gram sorbent, has been replaced by a concentration ratio. Replacing  $\Gamma$  by a concentration ratio was found to improve parameter estimation and is shown to be useful for data where the extent of adsorption is inferred from the difference between a reference concentration and the equilibrium value.

#### B. THEORETICAL CONSIDERATIONS

The difficulties in parametrizing isotherm equations are well known (see for example Dowd and Riggs, 1965; Harter, 1984; and Low and Batley, 1988). The pitfalls of linear regression may be illustrated by the Langmuir adsorption isotherm. For adsorption of a gas, the equation for the Langmuir isotherm can be written as:

$$\Gamma = \frac{c \Gamma_m \phi}{1 + c \phi} \quad \text{VI.1}$$

where  $\Gamma_m$  = the mass of gas required to form a monolayer per gram adsorbent (g/g)  
 $c$  = a dimensionless affinity constant related to the vapor binding energy  
 $\Gamma$  = the mass of gas adsorbed per gram adsorbent (g/g)  
 $\phi$  = adsorbate vapor concentration normalized by the concentration of the saturated vapor at the experimental temperature and pressure.

In this equation,  $\Gamma_m$  and  $c$  are parameters, whose values remain fixed, while the quantities  $\Gamma$  and  $\phi$  may vary as different cases are observed. The Langmuir model is appropriate if Equation VI.1 is satisfied by every observed pair  $\Gamma$  and  $\phi$ , when these are measured with perfect accuracy (i.e., without error).

The Langmuir equation has at least three linear transformations:

$$\Gamma = \Gamma_m - \frac{\Gamma}{c\phi} \quad \text{VI.2}$$

$$\frac{\phi}{\Gamma} = \frac{1}{c \Gamma_m} + \frac{\phi}{\Gamma_m} \quad \text{VI.3}$$

$$\frac{1}{\Gamma} = \frac{1}{c \phi \Gamma_m} + \frac{1}{\Gamma_m} \quad \text{VI.4}$$

All three of these equations are equivalent to VI.1, so all of the Equations VI.1 to VI.4 represent the Langmuir model.

In practice, there are two complications with use of Equations VI.1 through VI.4: the values of the Langmuir model parameters  $\Gamma_m$  and  $c$  are not known, and the pair  $\Gamma$  and  $\phi$  cannot be observed with perfect accuracy. Thus, the  $i^{\text{th}}$  experimental observation consists of a pair of values  $\Gamma_i$  and  $\phi_i$  that make Equations VI.1 to VI.4 hold approximately (but not exactly). Estimates of the Langmuir isotherm equation parameters  $\Gamma_m$  and  $c$  can be obtained from a regression analysis of a set of  $n$  observations, using one of the Equations VI.1 to VI.4. If the regression is based on one of the equations VI.2 to VI.4,  $\Gamma_m$  and  $c$  can be estimated from the slope and intercept of the best line drawn through the transformed data. Statistical research on this problem has focused on determining which transformation gives the most reliable parameter estimates. Some of the leading work in this area may be found in regard to the Michaelis-Menten equation in the field of enzyme kinetics (Dowd and Riggs, 1965), which has the same hyperbolic form as the Langmuir equation (Lineweaver and Burk, 1934).

Each linear transformation shown above suffers from an apparent drawback. The first two equations are deficient, as noted by Dowd and Riggs (1965), in that the same variable appears on both sides of the equation ( $\Gamma$  in Equation VI.2 and  $\phi$  in Equation VI.3), creating a correlation between the plotted variables. Equation VI.4 is also problematic because the reciprocal of  $\Gamma$  tends to emphasize small values, which are known with the least accuracy. Using synthetic data, Dowd and Riggs (1965) found that Equation VI.3 gave the most reliable parameter estimates for unweighted data and was far superior in this respect to the double reciprocal, or Lineweaver-Burk method (Equation VI.4) regardless of the error structure assumed. Paradoxically, Equation VI.4 created the line with the best visual appearance even though its ability to estimate the true parameters was poorest. In addition, it was found that the linear transformations having high correlation coefficients on the predictors and narrow confidence intervals for the parameters did not show superior ability to estimate parameter values correctly. In fact, linear transformations that highlighted outliers and resulted in broader 95 percent confidence intervals proved better at judging parameter estimates than transformations that obfuscated errors (Dowd and Riggs, 1965).

Harter (1984) showed, by use of examples, that both a linear appearance of a linearized Langmuir plot and a high value of the correlation coefficient were obtainable even when the data clearly did not fit the Langmuir model. This was apparent to the observer when the data was plotted without linearization along with the fitted curve. He concluded that linearity of a Langmuir plot derived from one of the Equations VI.2 - VI.4 did not sufficiently justify choosing the Langmuir model. He showed that several curves with different shapes could produce straight line plots using Equation VI.3 with high correlation coefficients. Small errors in the data could throw the prediction of the adsorption maximum off by as much as 50 percent. Harter concluded that, even when correlation coefficients are high, it is important to plot the data along with the fitted curve to check the fit. Absence of data in the high concentration range, precludes validation of the predicted adsorption maximum.

## **1. The Application of Regression Analysis**

Although the methodology for parameter estimation in regression models is highly developed, confusion may still result when regression analysis is applied to adsorption data, because of the nature of the adsorption experiments and the complexity of the isotherm equations. Important questions include which transformations of the isotherm equation and data are appropriate, how to avoid gross errors in parameter estimates due to extraneous variables, how to use weighted regression, how uncontrolled variations in experimental conditions can be handled, and whether the added complexities of approaches avoiding these difficulties are justified.

The derivation of the regression analysis equations involves four critical assumptions regarding the statistical properties of the data. Because departure from the assumptions inherent in regression models can lead to bias in parameter estimation and other difficulties, the four assumptions of regression are briefly reviewed.

Regression analysis relates the value of a response variable  $Y$  to the corresponding value of the predictor variable  $X$ . (The discussion here is restricted to regression with a single predictor variable, which suffices for the models considered in this report. It can be extended in straightforward fashion to the case of several predictor variables; see Draper and Smith, 1981.) The first assumption is that for every possible value of the predictor variable  $X$  there is a normal distribution of the dependent variable  $Y$ . The second assumption is that the values of  $X$  are known without error, or with negligible error. The third assumption is that the mean and the variance of the distribution of  $Y$  conditional on  $X$  are related to the value of  $X$  by known functional forms. It is often assumed that the variance of  $Y$  given  $X$  is the same for all  $X$ ; however, this standard assumption can be relaxed by using weighted least squares, which allows unequal variance problems to be recast as problems in which there are equal variances. The fourth assumption is that observations of the dependent variable  $Y$  are jointly independent (i.e., one measurement of  $Y_i$  does not affect any other).

To summarize, the four assumptions are:  $Y$  given  $X$  has a normal distribution; the values of  $X$  are measured without error; the distribution of  $Y$  given  $X$  has mean and variance of known form; and observations are independent. Point estimation of the parameters can be performed with less restrictive assumptions, e.g., dropping the requirement of normality and replacing independence of the  $Y$ 's by the weaker assumption that the  $Y$ 's are pairwise uncorrelated. However, the four assumptions as specified above are needed to obtain confidence intervals and test hypotheses. These assumptions are incorporated into the general model:

$$Y_i = \Gamma(X_i, \Gamma_m, c) + \varepsilon_i \quad \text{for } i=1,2, \dots, n, \quad \text{VI.5}$$

where:  $\Gamma(X_i, \Gamma_m, c)$  is the expectation function, of known mathematical form, evaluated at the value of the predictor variable,  $X_i$ ,

and  $\varepsilon_i$  is the random error term, which is normally distributed with mean zero and variance given by a function  $q(X_i, \Gamma_m, c)\sigma^2$  of known form. [If the variances of all observations are equal, the variance function reduces to  $\sigma^2$ .]

## 2. Plotting and Analyzing Adsorption Data

Since equations can often be written in many equivalent forms, e.g., Equations VI.1 - VI.4, a pressing problem is deciding which arrangement to use. Because of the first assumption of regression analysis, it is important to resolve the relationships between the variables. Some variables are controlled by the experimenter and some are not because they are the result of the phenomenon that the experimenter is attempting to study. Here we will refer to the experimental results as the response variable, or dependent variable. Similarly we will refer to what is varied in a controlled manner by the investigator as the predictor variable, or independent variable. Note that the same variable may be either a response variable or a predictor variable depending on the experimental design.

Wherever possible, the isotherm equation should be arranged, and plotted, so that the response variable is expressed as a function of the predictor variables and the parameters. When the experimenter ignores this suggestion, by plotting on a different set of axes, he or she inevitably introduces correlation between the x and y variables. The form of the isotherm equation plotted will depend on the experimental method chosen, to reflect the diverse roles of variables in different experimental protocols. With the experiments described here, equilibrium concentration is the response variable and sorbent mass is the predictor variable. Instead, if the experiment were performed by controlling the initial amount of sorbate and measuring the equilibrium concentration, under conditions where sorbent mass was held constant, then equilibrium concentration would be the response variable and initial concentration the predictor variable. The experimentally based rearrangement of the isotherm equation, once found, forms the basis of the analysis and requires no further manipulation before regression analysis (Kinniburgh 1986). This work focuses on the equations pertaining to a protocol where mass of adsorbent is chosen by the investigator and concentration is measured relative to a control level. This analysis may easily be adapted to other protocols.

The headspace analysis technique for vapor sorption will be used to illustrate the effect of experimental protocol on data analysis. Details of the headspace analysis technique, used here and by Peterson et al. (1988) to measure vapor sorption, are given in the materials and methods section.

In the headspace analysis technique the investigator starts by choosing a level of sample mass to weigh into an experimental bottle. Thus the sample mass in a bottle is the independent or assigned variable and will be referred to as the predictor. The equilibrium vapor concentration, estimated by gas chromatography, is then the dependent or measured variable, or the response variable, because it is a function of the mass of sample present. Sample mass is a good choice for the independent variable, because it may be measured

with a precise analytical balance ( $\pm 0.0002\text{g}$ ) and therefore best satisfies the assumption of regression analysis that the predictor variable is known without error. In the headspace procedure, as in many other experimental protocols, the amount of adsorption is not measured directly. Instead vapor concentration is measured and adsorption is inferred by mass balance from an initial vapor concentration. The amount of vapor adsorbed per gram, therefore, is viewed not as a variable in its own right, but as a quantity calculated from the values of other system variables.

If the choice of sample mass and gas concentration as predictor and response variables is ignored, one can construct the same plot as that used by experimenters where adsorption is measured directly (hereafter referred to as the conventional method), that is, a plot of the adsorption density (the mass of vapor adsorbed per gram of solid) versus equilibrium concentration. Analysis of headspace data by the conventional approach is complicated because both plotted variables are subject to errors, instead of solely the response variable as the regression model assumes. Excessive error propagation occurs with the conventional method since the errors occurring in the predictor and response variables are negatively correlated, shifting data points perpendicular to the true isotherm. For instance, a slightly larger estimate of the equilibrium concentration will produce a correspondingly lower estimate of adsorption. Conversely, a slightly lower estimate of concentration will result in a slightly greater estimate of adsorption. Because a unique distribution of Y for a given level of X is not guaranteed, and because errors occur in both the X and Y variables when the data are plotted as amount adsorbed per gram versus concentration, this approach is avoided here. An alternate equation form is proposed, which avoids the negative correlation of the variables and obeys the first regression assumption. The approach (derived below), which satisfies these criteria, is extended from the linear isotherm equation to the Langmuir and B.E.T. isotherms.

### 3. Adsorption Isotherms

Linearity of adsorption isotherms is commonly observed when relatively few molecules of adsorbate are present in the company of a large excess of adsorbent. This condition occurs at the low concentration range of the isotherm. Many experimental isotherms are linear over some range, especially those for partitioning of nonionic pollutants onto soils. Within the linear range of an isotherm, the relationship between the amount adsorbed and the concentration can be expressed via a single constant,  $K'_d$ .

In a typical vapor adsorption experiment the amount adsorbed is inferred by comparison of the equilibrium concentrations in a control (no adsorbent) vessel and a sample (with adsorbent) vessel. To the degree possible the same initial mass of adsorbate

is added to each vessel. Assuming a linear partitioning process, the equation for mass balance over controls and samples can be equated as:

$$V_B C_B = \Gamma_i M_i + V_i C_i \quad \text{VI.6}$$

where:  $V_B$  = the average volume of the control bottles ( $\text{cm}^3$ )  
 $C_B$  = the average concentration of the controls ( $\text{g}/\text{cm}^3$ ),  
 $M_i$  = the sample mass (g),  
 $V_i$  = the volume of the sample bottle ( $\text{cm}^3$ ),  
 $C_i$  = the concentration measured for the sample bottle ( $\text{g}/\text{cm}^3$ ).

If vapor density is measured by gas chromatography, then concentration is proportional to G.C. signal area:

$$C_B = \Phi_B z/B \text{ and } C_i = \Phi_i z/B, \quad \text{VI.7}$$

where:

$\Phi_B$  = the average signal area of the control bottles,  
 $\Phi_i$  = the measured signal area of the sample bottle,  
 $z$  = the density of the saturated vapor ( $\text{g}/\text{cm}^3$ )  
 $B$  = the signal area equivalent of the saturated vapor or  $\Phi_B/(V_G/V_B)$ , where  
 $(V_G/V_B)$  = a volume dilution factor, and  
 $V_G$  = the volume of the source syringe ( $\text{cm}^3$ ).

Approximation of gas concentration by G.C. signal area allows Equation VI.6 to be written as:

$$\frac{V_B \Phi_B z}{B} = \Gamma_i M_i + \frac{V_i \Phi_i z}{B} \quad \text{VI.8}$$

When a direct relationship exists between the amount adsorbed, for a given mass of adsorbent,  $M_i$ , and the relative concentration of the sorbate,  $\phi$ , where  $\Phi_i/B = \phi_i$ , a linear isotherm (Equation VI.9) may be substituted for  $\Gamma_i$  in Equation VI.8.

$$\Gamma_i = K \Phi_i/B = K'_d z \Phi_i/B \quad \text{VI.9}$$

where  $K$  is the dimensionless parameter to be fitted,  
and  $K'_d$  has dimensions of  $\text{cm}^3/\text{g}$ .

Equation VI.8 can then be arranged to put the dependent variable ( $\Phi_i$ ) that is measured on the left hand side, keeping the independent variable ( $M_i$ ) and the parameter to be fitted ( $K$ ) on the right:

$$\frac{V_B \Phi_B}{V_i \Phi_i} = K \frac{M_i}{V_i z} + 1 \quad \text{VI.10}$$

Equation VI.10 may also be derived by observing that  $\Gamma_i$  may be expressed as:

$$\Gamma_i = \frac{z(V_B \Phi_B - V_i \Phi_i)}{M_i B} \quad \text{VI.11}$$

and substituting into Equation VI.9.

Defining for observation  $i$  ( $i=1,2, \dots, n$ ),

$$Y_i = \frac{V_B \Phi_B}{V_i \Phi_i} \quad \text{and} \quad X_i = \frac{M_i}{V_i z} \quad \text{VI.12}$$

transforms Equation VI.10 into  $Y_i = KX_i + 1$ . If the assumptions of the model VI.5 are satisfied at least approximately [the linear isotherm is the simplest form of Equation VI.5 where  $K = f(\Gamma_m c)$ ], this model can be used to analyze the observations ( $X_i, Y_i$ ). The expectation function ( $KX_i + 1$ ) gives the linear model:

$$Y_i = K X_i + 1 + \epsilon_i \quad \text{for } i=1, 2, \dots, n, \quad \text{VI.13}$$

with  $\epsilon_i$  representing an error term, required whenever observations of  $Y_i$  are not exact. As will be seen below, for the experimental data presented here, an exponential function best explained the squared deviation of  $Y_i$  from the expected value of  $Y_i$  ( $\hat{Y}_i$ ). Details of the model used for the variance of the observations, and the manner in which the variances were examined, are given in the data analysis and results and discussion sections.

It is possible, of course, that no choice of the variance function will make a linear model suitable for a given set of observations. This will be the case if the assumption of a linear partitioning process is not satisfied. In this situation, an alternate model may be more consistent than the linear model (VI.9) with the observations. Two nonlinear models, the Langmuir and the B.E.T. equations, will now be treated.

#### 4. Fitting Nonlinear Models: The Langmuir and B.E.T. Equations

The model connecting a response variable  $Y$  and a predictor variable  $X$  can lead to a regression analysis of varying degrees of complexity. The simplest case is the linear regression model of Equation (VI.9). A more general case is the regression model of Equation (VI.5), which does not require the function  $\Gamma$  to be linear in the parameters.

It is possible that  $X$  and  $Y$  satisfy a physical law that does not allow an explicit solution for  $Y$  as a function of  $X$  and the parameters. If  $X$  and  $Y$  could both be measured without error, all observed  $X, Y$  pairs would satisfy a relationship of the form:

$$f(X, Y; \{\theta\}) = 0, \quad \text{VI.14}$$

where  $\{\theta\}$  denotes the set of parameters in the model. The Redlich-Peterson isotherm is an example of an equation which satisfy Equation VI.14 but not Equation VI.5. The crucial point is that regression methods can still be applied in spite of the inability to solve for  $Y$  explicitly as was done in Equation VI.5.

For specified values of  $X$  and the parameter set  $\{\theta\}$ , the value(s) of  $Y$  that satisfy Equation VI.14 can be determined. Thus, even if  $Y$  can not be explicitly solved for,  $f_Y(X; \{\theta\})$  can be defined implicitly as the value of  $Y$  that satisfies VI.14 for a given  $X$  and  $\{\theta\}$ . To reflect the fact that  $Y$  is observed with error, an error term is incorporated in the relation between  $Y$  and  $f_Y(X; \{\theta\})$ .

$$Y = f_Y(X; \{\theta\}) + \epsilon. \quad \text{VI.15}$$

Although  $f_Y$  is an implicit rather than an explicit function, Equation VI.15 can be solved numerically as needed and the results used to obtain a regression analysis of the data. Even when it is possible to solve explicitly for  $Y$ , such as with the Langmuir isotherm, it may still be advantageous to use an implicit solution to a nonlinear adsorption isotherm to avoid possible colinearity in the coefficient matrix for the parameters which may occur under nonlinear regression on Equation VI.1.

##### a. The Langmuir Equation

Langmuir derived a model isotherm (VI.1) for an adsorption process in which adsorption proceeds to a monomolecular film because at most one molecule is allowed to occupy a given adsorption site. Brunauer, Emmet, and Teller (1938) extended Langmuir's equation to allow adsorbate molecules to form multiple layers by relaxing this assumption.

The B.E.T. equation, as their relationship came to be known, assumes that a dynamic equilibrium (similar to that described by the Langmuir equation) explains adsorption for each successive layer. Both Langmuir and Brunauer, Emmet and Teller originally developed their isotherms to describe the adsorption of gases by solids.

The data analysis for the Langmuir and B.E.T. model equations is similar to that for the linear model and involves the same variables X and Y (defined in Equation VI.12). As was the case with the linear isotherms, it is dimensionally expedient to normalize the measured adsorbate concentration (or G.C. signal area) by the saturated vapor pressure concentration (i.e., the G.C. signal area expected at that concentration).

$$\Gamma = \frac{\Gamma_m c \frac{\Phi_i}{B}}{1 + c \frac{\Phi_i}{B}} \quad \text{VI.16}$$

where  $\Gamma_m$  = mass of adsorbate corresponding to monolayer coverage per gram of adsorbent,  
 $c$  = the Langmuir affinity parameter,  
 $\Gamma$  = mass of vapor adsorbed per gram adsorbent.

Substitution of Equations VI.12 and VI.16 into Equation VI.8 yields:

$$Y_i = X_i \frac{c \Gamma_m}{1 + c Z_i Y_i^{-1}} + 1 + \epsilon_i \quad \text{for } i = 1, 2, \dots, n \quad \text{VI.17}$$

where  $Z_i = V_G/V_i$ ,  $V_G$  is the volume of the source syringe ( $\text{cm}^3$ ), and the subscript  $i$  refers to individual observations.

For the Langmuir equation an implicit relationship, as in Equation VI.15, between X and Y was chosen as the basis for analysis. An error term  $\epsilon_i$  is incorporated into Equation VI.17 to reflect the random variability in  $Y_i$ . This reflects the fact that several observations with identical values of  $X_i$  will result in distinct values of  $Y_i$ . Regression analysis can still be performed, despite the occurrence of the  $Y_i^{-1}$  on the right-hand side of Equation VI.17, when a numerical procedure is used to calculate values for  $\Gamma_m$  and  $c$  that minimize the residual sum of squares (RSS). The details of fitting the parameters and incorporating the error in the observations is developed in the data analysis section.

b. The B.E.T. Equation

The form of the B.E.T. isotherm used here is:

$$\Gamma = \frac{\Gamma_m c \frac{\Phi_i}{B}}{\left(1 - \frac{\Phi_i}{B}\right) \left\{1 + (c-1) \frac{\Phi_i}{B}\right\}} \text{ for } i = 1, 2, \dots, n \quad \text{VI.18}$$

where terms are defined as in Equation VI.16. This can be treated as the Langmuir equation was previously, to form an equation in X and Y of Equation VI.12, by substituting for  $\Gamma$  in Equation VI.18 by Equation VI.11:

$$Y_i = X_i \frac{\Gamma_m c}{\left[1 - \frac{Z_i}{Y_i}\right] \left[1 + (c-1) \left(\frac{Z_i}{Y_i}\right)\right]} + 1 + \epsilon_i \quad \text{VI.19}$$

The equation used for the B.E.T. isotherm, as was the case with the Langmuir isotherm, is an implicit function of  $Y_i$  in terms of  $X_i$  and the parameters when the extent of adsorption is not independently measured. Equation (VI.19), like Equations VI.13 and VI.17, includes an error term  $\epsilon_i$  to reflect the random variability in  $Y_i$ . Regression analysis may proceed despite the implicit nature of the function for  $Y_i$  as described in the data analysis section which follows.

## 5. Weighting Data for Regression

Having identified the dependent variable in the system, it is appropriate to examine the issue of equality of variances. Frequently some observations in a regression analysis are "less reliable" than others. The variances of the observations will hence be unequal, and weights corresponding to these variances must be assigned to the observations. Analysis of these weighted observations involves transforming them to a set of variables to which we can apply the usual (unweighted) regression analysis. A detailed treatment of weighted least squares regression is given by Draper and Smith (1981).

To understand the variance structure of the observations, it is useful to identify potential sources of error. For instance, in this work, in which vapor concentration was measured, error was likely attributable to the difficulty of measuring low concentrations on the G.C. relative to the blanks, which were used to set the signal range. This error was compounded by the division operation (i.e., calculation of  $Y_i$ ), especially as the divisor

approached zero. In general there is no method of knowing *a priori* how the variance of a random variable will behave over its entire range. It is good practice, therefore, to conduct replicate experiments so that errors may be quantified. In absence of a sizable pool of replicates, it may be appropriate to assume a functional relationship for the variance in the response variable based on the magnitude of the response or predictor variable. Common assumptions of this type are discussed by Kinniburgh (1986). The model used for the variance of vapor sorption data is described in the Results and Discussion section below.

## 6. Fitting the Intercept

The model forms of the linear, Langmuir, and B.E.T. isotherms (Equations VI.13, VI.17, and VI.19, respectively) share a common intercept of 1.0. The fitting routine used here forced the linear model through the theoretical value of the intercept. Poor fit indicated that an alternate model should be chosen. Forcing the intercept through the theoretical value was necessary to make different forms of the linear model equivalent. Intercepts of the nonlinear models were not forced.

## C. DATA ANALYSIS

### 1. The Langmuir Equation

Although equations VI.17 and VI.19 may be solved explicitly, an implicit solution may also be employed. We elected to solve the equation implicitly for  $Y_i$ , in attempt to improve parametrization by nonlinear regression. The values of the parameters  $\Gamma_m$  and  $c$  that minimize the RSS were found by using a numerical minimization algorithm. The numerical procedure was simplified by the fact that the Langmuir equation is linear in the parameter  $\Gamma_m$ . This made it possible to treat only  $c$  iteratively, rather than the pair  $(c, \Gamma_m)$ . For any given value of  $c$ , the usual linear regression methods can be applied to find the value of  $\Gamma_m$  that minimizes the RSS for the data and the specified value of  $c$ . We proceeded by iterating on  $c$ , finding for each value of  $c$  the minimizing value of  $\Gamma_m$  and the resulting minimum RSS. In summary, the residual sum of squares was the objective function minimized and  $c$  was the decision variable.

Minimization was performed by Golden Section search which is described in detail by Press et al. (1986), and Wagner (1969). This method begins with the value of  $c$  bracketed in an initial interval  $[a,b]$ , specified by the user. With each iteration, the function is evaluated at two interior points,  $C_1$  and  $C_2$ . The points  $C_1$  and  $C_2$  are chosen according to golden ratio to optimize convergence. They divide the bracketing interval into three sections. By looking at the value of the objective function at the interior points  $C_1$  and  $C_2$ , one of the end sections of the bracketing interval is discarded. The process continues until

the interval around the  $c$  that minimizes the RSS is tolerably small, as described by the following algorithm:

```

set   C0=a
      C3=b
      C1=0.618a + 0.392b           and           RSS1=RSS(C1)
      C2=0.392a + 0.618b           and           RSS2=RSS(C2)
Do the following loop while the absolute value of (C3-C0) is greater than the tolerance:
  If RSS2 is less than RSS1 then
    C0=C1;           and           RSS0=RSS1
    C1=0.392C0 + 0.618C3;   and           RSS1=RSS(C1)
    C2=0.618*C0+0.392*C3;   and           RSS2=RSS(C2)
  else
    C3=C2           and           RSS3=RSS2
    C2=0.392*C3 + 0.618*C0   and           RSS2=RSS(C2)
    C1=0.618*C3+0.392*C0   and           RSS1=RSS(C1)
End

```

For each specified value of  $c$ , the value of  $\Gamma_m$  minimizing RSS (Equation VI.29) was determined from the linear regression equation with no intercept:

$$U_i = X_i \Gamma_m + \delta_i \quad \text{for } i = 1, 2, \quad \text{VI.20}$$

where  $\delta_i$  is a random error term and  $U_i$  is defined as the function of  $Y_i$ :

$$U_i = (1/c) [Y_i + c Z_i(1 - Y_i^{-1}) - 1] \quad \text{VI.21}$$

Equation VI.21 is obtained from Equation VI.17 by moving all terms involving  $Y_i$  to the left-hand side while keeping all occurrences of  $X_i$  and  $\Gamma_m$  on the right-hand side. We assume in VI.20 that the  $n$  observations are independent and that the random error term  $\delta_i$  has a normal distribution with mean zero and variance given by:

$$\text{Var}(U_i) = \text{Var} \left\{ (1/c) [Y_i + c Z_i(1 - Y_i^{-1}) - 1] \right\} \quad \text{VI.22}$$

The variance can be approximated by the standard result, based on a first-order Taylor series expansion assuming the variance of  $Z_i$  is negligible (see Kendall and Stuart, 1969):

$$\text{Var}(g(Y_i)) \approx (d g(Y)/dY)^2 |_{Y=Y_i} \text{Var}(Y_i). \quad \text{VI.23}$$

Applying this to the function  $U_i = g(Y_i) = c^{-1}[Y_i + c Z_i(1 - Y_i^{-1}) - 1]$  gives:

$\text{Var}(U_i) \approx d_i \text{Var}(Y_i)$ , where  $d_i$  is defined by:

$$d_i = c^{-2} (1 + (c Z_i) Y_i^{-2})^2. \quad \text{VI.24}$$

The technique of weighted least squares may now be used to estimate the parameter  $\Gamma_m$  in VI.20, where the weight of each observation is proportional to the reciprocal of its variance. For a discussion of weighted least squares regression see Draper and Smith, (1981).

Examination of the data obtained in the vapor sorption experiments reported here suggested that the variance of  $Y_i$  can be modeled as a function of  $\hat{Y}_i$ , the expected value of  $Y$  given the value of  $X_i$ , where  $h$  is approximately equal to the variance of the blanks and  $Q$  is a power constant. The choice of this particular function for modeling the variance will be justified in the discussion of results.

$$\text{Var}(Y_i) = h \hat{Y}_i^Q \quad \text{VI.25}$$

The weights used with Equation (VI.20) are derived from Equation (VI.25) as in Draper and Smith (1981). The variance of  $U_i$  is calculated from the variance of  $Y_i$  as:

$$\text{Var}(U_i) = \frac{\sigma^2}{w'_i} \approx d_i h Y_i^Q \quad \text{VI.26}$$

yielding the weights applicable to estimation of the parameter  $\Gamma_m$  of Equation VI.20,

$$w'_i = \frac{\sigma^2}{h Y_i^Q}. \quad \text{VI.27}$$

Then  $\hat{\Gamma}_m$  is given by weighted linear regression (Draper and Smith, 1981):

$$\hat{\Gamma}_m = \frac{\sum_{i=1}^n w'_i X_i U_i}{\sum_{i=1}^n w'_i X_i^2} = \frac{\sum_{i=1}^n \frac{X_i U_i}{d_i Y_i^Q}}{\sum_{i=1}^n \frac{X_i^2}{d_i Y_i^Q}} \quad \text{VI.28}$$

This is the  $\hat{\Gamma}_m$  that minimizes the weighted RSS for the data  $(X_i, Y_i)$ ,  $i=1,2, \dots, n$ , and the given  $c$ . The RSS is given by:

$$\text{RSS}(c, \hat{\Gamma}_m) = \sum_{i=1}^n w_i (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n w_i R_i^2 \quad \text{VI.29}$$

where residuals  $R_i = Y_i - \hat{Y}_i$  are calculated using Equation VI.34 given below. The variance of  $Y_i$  is given by Equation VI.25, and setting this equal to  $\sigma^2/w_i$  (Draper and Smith, 1981) gives:

$$w_i = \frac{\sigma^2}{h Y_i^Q} \quad \text{VI.30}$$

Therefore:

$$\text{RSS}(\hat{c}, \hat{\Gamma}_m) = \sum_{i=1}^n w_i (Y_i - \hat{Y}_i)^2 \quad \text{VI.31}$$

The residuals are calculated as the difference between  $Y_i$  and  $\hat{Y}_i$  where  $\hat{Y}_i$  is the value for  $Y$  predicted using the estimated values of  $\Gamma_m$  and  $c$ , denoted by  $\hat{\Gamma}_m$  and  $\hat{c}$ . The positive root to the quadratic equation:

$$Y_i^2 + (\hat{c} Z_i - X_i \hat{\Gamma}_m \hat{c} - 1) Y_i - \hat{c} Z_i = 0 \quad \text{VI.33}$$

gives the predicted value for  $Y$ ,  $\hat{Y}_i$ , of Equation VI.17, as:

$$\hat{Y}_i = \frac{1}{2} \left[ \hat{c} \hat{\Gamma}_m X_i - \hat{c} Z_i + 1 + \sqrt{\hat{c}^2 Z_i^2 - 2 \hat{c} \hat{\Gamma}_m X_i Z_i + 2 \hat{c} \hat{\Gamma}_m X_i + 2 \hat{c} Z_i + (\hat{c} \hat{\Gamma}_m X_i)^2 + 1} \right] \quad \text{VI.34}$$

## 2. The B.E.T. Equation

Data analysis for the B.E.T. equation was similar to the Langmuir equation except that solving for  $\hat{Y}_i$  resulted in a cubic expression that was evaluated as described by Abramowitz and Stegun (1972). In the case of the vapor sorption data the determinant was negative, implying three real roots. The first root was the desired solution, the second root being negative and the third less than 1.0.

Like the Langmuir equation, the B.E.T. equation is linear in the parameter  $\Gamma_m$ . A similar numerical analysis was therefore applied. A value for  $c$  was chosen and weighted linear regression (the left-hand portion of Equation VI.28) was used to solve for  $\Gamma_m$ . The regression equation resulting from (VI.19) was

$$T_i = X_i \Gamma_m + \delta_i \quad \text{for } i=1, 2, 3, \dots, n \quad \text{VI.35}$$

where  $\delta_i$  is a random error term and  $T_i$  is defined as:

$$T_i = g(Y_i) = (1/c)(Y_i + a_1 + a_2 Y_i^{-1} + a_3 Y_i^{-2}) \quad \text{for } i=1, 2, 3, \dots, n \quad \text{VI.36}$$

where

$$a_1 = (c - 2)Z_i - 1, \quad \text{VI.37}$$

$$a_2 = -(c - 2)Z_i - (c - 1)Z_i^2, \quad \text{VI.38}$$

$$a_3 = (c - 1)Z_i^2, \quad \text{VI.39}$$

This is obtained from Equation VI.19 by moving all terms involving  $Y_i$  to the left-hand side while keeping all occurrences of  $X_i$  and  $\Gamma_m$  on the right-hand side. As with the Langmuir equation, we assume that the  $n$  observations are independent and that the random error term  $\delta_i$  has a normal distribution with mean zero and variance given by:

$$\text{Var}(T_i) \approx \text{Var}\left\{\frac{1}{c}[Y_i + a_1 + a_2 Y_i^{-1} + a_3 Y_i^{-2}]\right\}. \quad \text{VI.40}$$

The analysis proceeds as before to estimate the variance of  $T_i$  from the variance of  $Y_i$  using a first-order Taylor series approximation:

$$\text{Var}(T_i) = g_i b Y_i^Q \quad \text{VI.41}$$

where:

$$g_i = \frac{1}{c^2} [1 - a_2 Y_i^{-2} - 2a_3 Y_i^{-3}]^2. \quad \text{VI.42}$$

Equations VI.41 and VI.42 are used to calculate the weights for the regression model for the B.E.T. Equation (VI.35) using an equation similar to VI.19.

The same method is used to calculate the residual sum of squares, including the same weights as in the case of the Langmuir equation. The RSS minimizing value of  $c$  was likewise found by golden section. Once predictions of  $\Gamma_m$  and  $c$  were obtained,  $\hat{Y}_i$  was the solution of:

$$\hat{Y}_i^3 + a_1 \hat{Y}_i^2 + a_2 \hat{Y}_i + a_3 = 0 \quad \text{for } i=1, 2, \dots, n. \quad \text{VI.43}$$

The fitting of both the Langmuir and the B.E.T. isotherms was done in FORTRAN on the Cornell IBM 4381 mainframe computer.

### 3. Results

Several approaches to data analysis were compared for 25 synthetic data sets for the linear and Langmuir isotherms generated with the SAS statistical computing program. Synthetic data allowed parameter estimates obtained in each analysis to be compared to the exact or "true" values. A value of one was chosen for the slope of the linear model and a value of eleven for the parameter  $c$  in the Langmuir model with  $\Gamma_m$  equal to one eleventh. Sample mass, assumed to be assigned without error, was allowed to have seven values ranging from five to thirty-five. Two replicates were used at each level of mass. Given the mass, the corresponding value of  $\Phi$  was calculated for each of two models, the linear and the Langmuir.  $\Phi_B$  was set equal to one. Two kinds of error structures were assumed: a) error of constant magnitude and b) constant percentage error. Error of constant magnitude was modelled by adding to  $\Phi_i$  a random variable with mean equal to zero and  $\sigma = 0.0005$ . Constant percentage error was incorporated by multiplying  $\Phi_i$  by a random variable with mean equal to one and  $\sigma = 0.03$ . For the linear model, two forms of linearization were compared for each error structure, Equations VI.9 and VI.13. The Langmuir model was fit to the synthetic Langmuir data by two linearizations, Equations (VI.3) and (VI.4), by

nonlinear regression on Equation VI.1, and by the Golden Section numerical approach on Equation VI.17. Weighting was used in all cases as the reciprocal of the variance given in Equation VI.23. Parameterization methods were compared using the computed sum of squared normalized differences between the true values and the estimates. Methods were tested for systematic bias with frequency diagrams (Figures VI.1-VI.8).

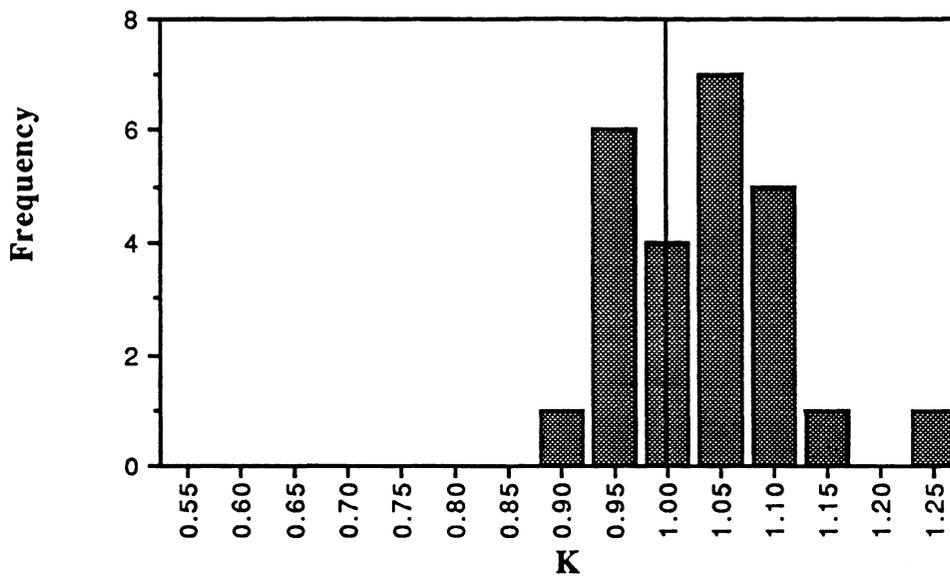


Figure VI.1. Frequency Distribution using Equation VI.9 on Data with Relative Error.

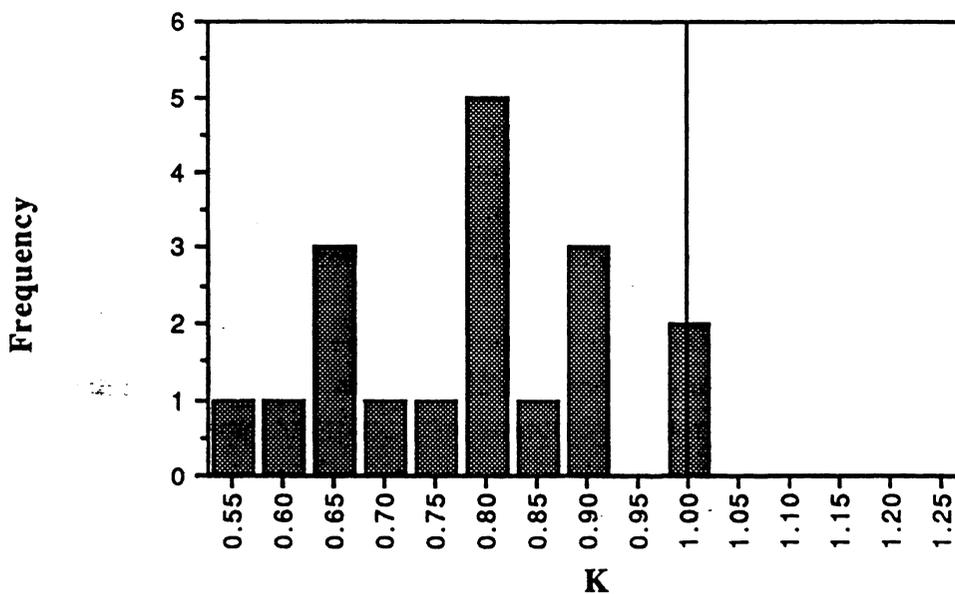


Figure VI.2. Frequency Distribution using Equation VI.9 on Data with Absolute Error.

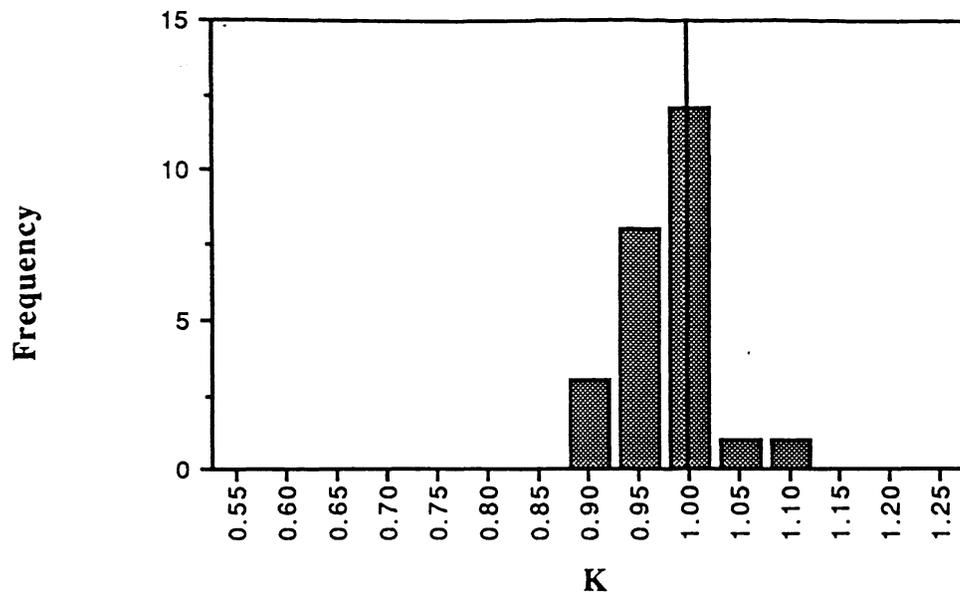


Figure VI.3. Frequency Distribution for K using Equation VI.13 on Data with Relative Error.

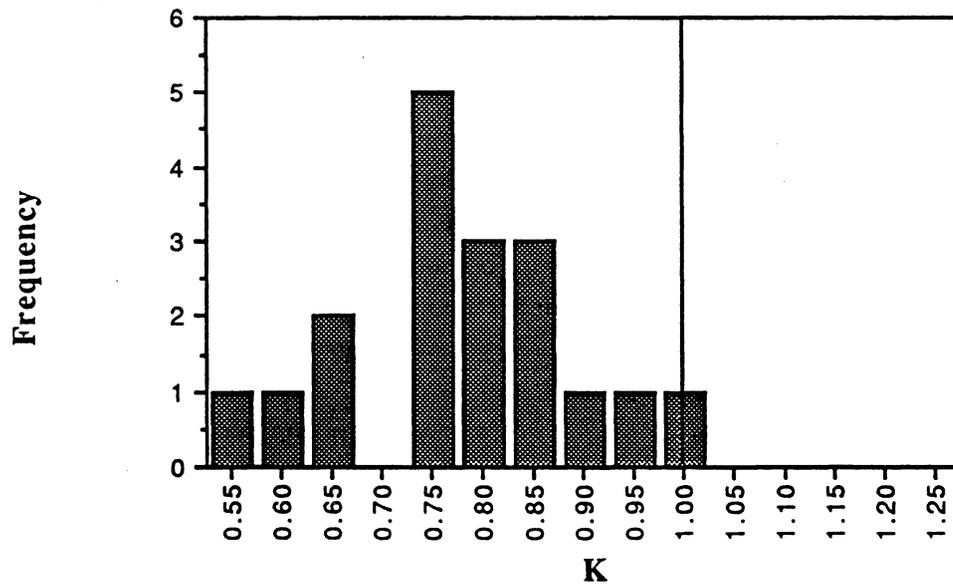


Figure VI.4. Frequency Distribution for K using Equation VI.13 on Data with Absolute Error.

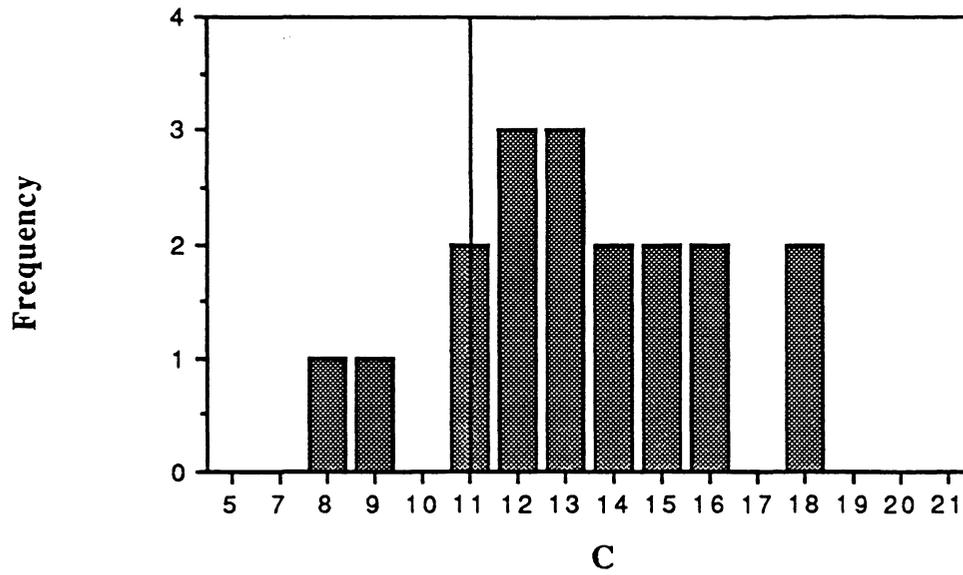


Figure VI.5. Frequency Distribution for c using Equation VI.3 on Data with Relative Error.

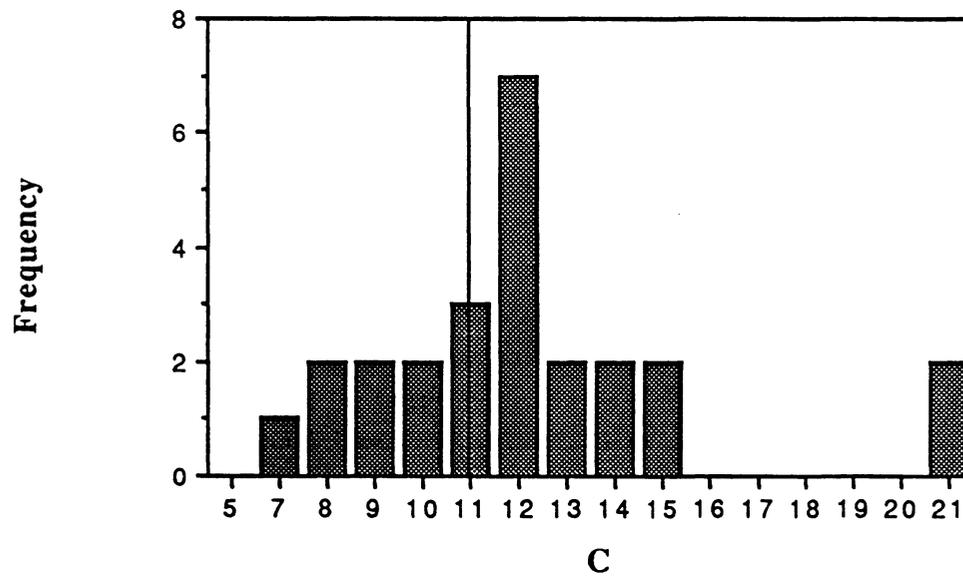


Figure VI.6. Frequency Distribution for c using Equation VI.3 on Data with Absolute Error.

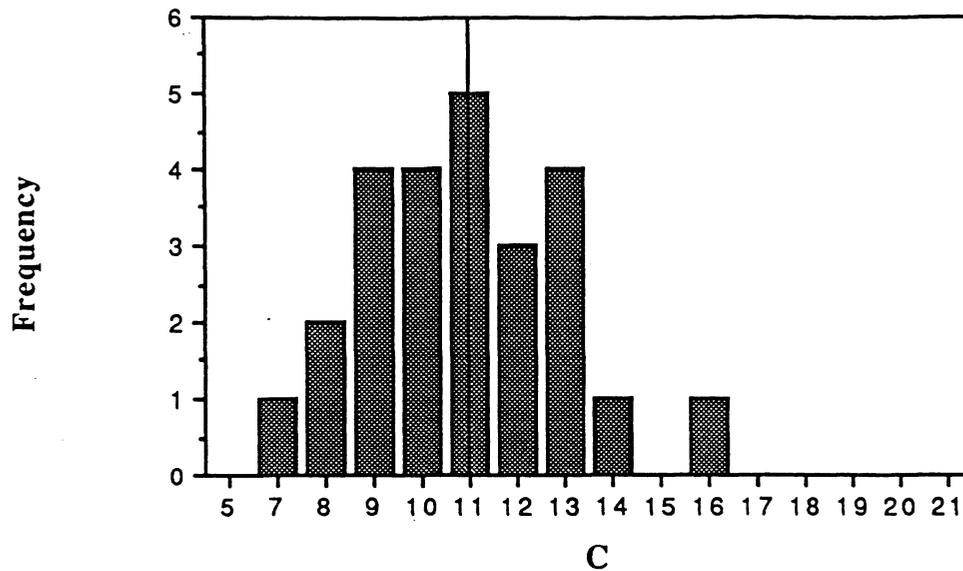


Figure VI.7. Frequency Distribution for  $c$  using Equation VI.17 on Data with Relative Error.

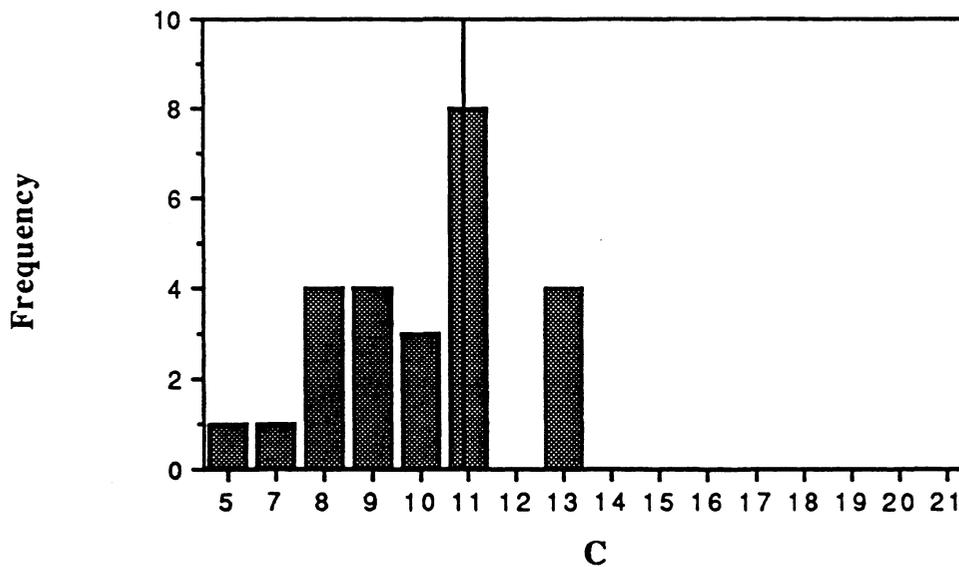


Figure VI.8. Frequency Distribution for  $c$  using Equation VI.17 on Data with Absolute Error.

Results of the data analysis comparison study are shown in Table VI.1. For both the constant magnitude and constant percentage error synthetic data sets, Equation VI.13, on the average, gave better estimates than Equation VI.9, although for the constant magnitude error the two methods performed virtually identically (due to the weighting). Equation VI.3 outperformed the method of nonlinear regression on Equation VI.1. The method of Golden Section was superior to both nonlinear regression and linearization by Equation VI.3. As found by Dowd and Riggs (1964), Equation VI.3 did better than Equation VI.4. The frequency diagrams illustrate that data with constant magnitude error tends to lead to underprediction of the slope of the linear equation. Considerably less bias was observed in the estimate of parameters from data with constant percentage error. Equation VI.3 for the Langmuir isotherm tended to overestimate the value of the parameter  $c$  as opposed to Golden Section which was more likely to underestimate it. Results for  $\Gamma_m$ , not shown, were negatively correlated with  $c$ . [Also not shown are the frequency diagrams for nonlinear regression (Equation VI.1) and linearization of the Langmuir equation by Equation VI.4 because both methods gave inferior results to those which are shown.] The methods for fitting the B.E.T. equation could be compared in a similar manner. Since there is no reason to believe that there would be a difference, this comparison is not provided here. In conclusion, the methods advocated in this work are proven for their ability to estimate parameters of synthetic data sets. Experimental data (see below) show that, for the Langmuir and B.E.T. equations, Golden Section was the most practical approach.

TABLE VI.1 SUMS OF SQUARED NORMALIZED ERRORS IN PREDICTIONS

	<u>Constant Magnitude error</u>	<u>Constant percentage error</u>
Linear Eqn (VI.9)	1.13	0.17
Linear Eqn. (VI.13)	1.11	0.064
Langmuir Eqn.(VI.3)	1.89	2.84
Langmuir Eqn.(VI.4)	28900	47.9
Nonlinear Regression Eqn. (VI.1)	4040	123
Golden Section Eqn. (VI.17)	1.14	1.80

#### D. MATERIALS AND METHODS

Six organic vapors were chosen to represent two classes of volatile pollutants found in aquifers: chlorohydrocarbon solvents and degreasers, and fuel hydrocarbons. Sorption of each vapor was followed in a separate experiment. The experimental compounds included chloroform (CF), methylene chloride (DCM), and 1,1,1-trichloroethane (1,1,1-TCA) from Fischer Scientific Co. perchloroethylene (PCE) and 1,1-dichloroethylene (1,1-DCE) from Aldrich Chemical Co. and toluene from Malinkrodt Chemical Co. All compounds were certified reagent grade.

Chromatographic adsorption alumina, 80-200 mesh size, (Fischer Scientific) was chosen for use as a sorbent because of its large surface area, well-characterized particle size distribution, and known chemical composition (Table VI.2). In some experiments the alumina was coated with humic acid (Aldrich Chemical Co.) using the procedure of Garbarini and Lion (1985). The purpose of the humic acid coating was to provide a surrogate organic phase for natural soil organic matter. Specific surface area was measured by B.E.T. nitrogen adsorption. Particle density was measured by the pycnometer method (ASTM, 1958). Pore size distributions were obtained by mercury porosimetry and from the nitrogen desorption curve (Lowell and Shields, 1984). Carbon content was determined by wet combustion with  $\text{Cr}_2\text{O}_7^{2-}$  (Allison, 1960).

TABLE VI.2 PROPERTIES OF SOLIDS INVESTIGATED

<u>Property</u>	<u>humic acid</u>	
	<u>coated alumina</u>	<u>uncoated alumina</u>
$\text{N}_2$ surface area ( $\text{m}^2/\text{g}$ )	189.3	143.2
mean pore size, $\text{N}_2$ ( $\text{\AA}$ )	20	20-30
organic carbon (percent)	0.45	0.02
particle density ( $\text{g}/\text{cm}^3$ )	2.57	2.98

In the headspace analysis technique, used here and by Peterson et al. (1988), the sorbent is exposed to a controlled amount of sorbate vapor in a vial until equilibrium is established. The equilibrated vapor in the headspace was analyzed by gas chromatography (Hewlett Packard Model 5890) using a flame ionization detector. The column was 100/180 mesh Supelcoport<sup>®</sup> coated with 20 percent SP-2100 and 0.1 percent Carbowax 1500. The signal areas obtained were compared to the average signal area measured for the controls (no sorbent). Masses of sorbent in the vials, in the experiments on dry alumina, were varied at five levels of three replicates each ranging typically between 0.05 and 1.0 grams.

Vapors were transferred to and from the vials with valved gas-tight syringes (Precision Sampling). One mL of saturated ( $25^\circ\text{C}$ ) vapor was used for all vapors except PCE (2.0 mL) and p-xylene (5.0 mL). The source for vapor was a 150 mL vessel containing roughly ten mL of liquid product in equilibrium with the airspace. Here as in the sample vessel, seal was provided by Teflon<sup>®</sup>-backed rubber septa (Supelco Chromatography Supply) with tear-away aluminum crimp seals. Five source bottles were used in rotation to ensure a constant source vapor concentration. In the experimental protocol, vapor was delivered to the open sample vessel with the syringe needle placed deep within the bottle. Septum integrity was preserved by leaving the cap ajar during vapor delivery. Incubation

periods at 25°C ranged between 12 and 24 hours, sufficient to allow equilibration based on preliminary experiments.

The experiment on p-xylene employed seventy-seven replicates and was used to define the structure of the variance associated with  $Y_i$ . Coated alumina, used in this study, was moistened by mechanical incorporation of distilled water to a moisture content of roughly 20 percent. Masses of moist coated alumina varied by eight levels (0.25, 0.5, 1.0, 2.0, 3.0, 5.0, 8.0, and 10.0 grams). Each of the eight mass levels was replicated in seven blocks. To each block was added three controls to form a total of eleven replicates per block. The replicates in each block were randomized.

## E. RESULTS AND DISCUSSION

### 1. Variance Model and Weighting

Replication within the sorption experiment on p-xylene allowed the error associated with the measurement of  $Y$  to be examined. Differences in moisture content, an important variable (Chiou et. al., 1985 and Chiou et. al., 1988) occurred between the blocks (Table VI.3). Such differences were incorporated into the sorption model using the following relationship:

$$Y_i = X_i (b_1 G_i + b_2) K + \epsilon_i \quad \text{for } i=1,2, \dots, n \quad \text{VI.44}$$

where  $G_i$  is the percent moisture content of sample  $i$  (g/g) and  $b_1$  and  $b_2$  are parameters and  $\epsilon_i$  is the error of sample  $i$ . The average of the term in parentheses is equal to one. This relationship allowed  $K$  to be computed for any moisture content, within the range of those investigated, so that moisture content differences could be included in the analysis. For p-xylene this relationship was:

$$Y_i = X_i (0.0130 - 0.0005653 G_i) K \quad R^2=0.92. \quad \text{VI.45}$$

TABLE VI.3 RESULTS OF LEAST SIGNIFICANT DIFFERENCE (LSD) ANALYSIS

<u>Grouping*</u>	<u>Mean</u>	<u>Block</u>
A	20.96	1
B	20.55	2
B C	20.46	3
B C	20.34	4
C	20.13	5
D	19.70	6
E	19.18	7

\*Mean moisture contents with the same letter are not significantly different at 95 percent.

The variance in Y is shown in Figure VI.9. The graph suggests that a log linear relationship of the form of Equation VI.25 may be used to describe the variance of Y. [Data points at  $\log(Y) = 1$  constitute blanks and are not well fit by the model.] The moisture content correction was used to adjust observed values of Y before fitting to Equation VI.25 (Figure VI.9). Incorporation of moisture content effects reduced the slope, Q, from  $3.104 \pm 0.111$  to  $2.621 \pm 0.159$  when all moisture contents were corrected to 20.4 percent. The variance structure observed in the p-xylene experiment was compared to that observed in the experiments on the six organic vapors where Q was  $2.621 \pm 0.159$  (see Figure VI.9.) Because values of Q were the same at the 95 percent confidence level, it was concluded that a single value of Q could be used throughout. Thus for all data analysis a value of  $Q = 2.687 \pm 0.017$  was used. This value was based on a compilation of eight experiments, (including the p-xylene data) including a total of 57 groups of three or more replicates of the level of sample mass. The relationship explained over 92 percent of the variation in the variance. It is assumed that this exponential relationship is adequate to explain the variance for the purpose of data weighting in regression. Experiments in the variance study included those on coated as well as uncoated alumina, dry as well as moist samples, chlorohydrocarbons as well as hydrocarbons. The relationship, which employs a single power constant Q and the variance of the controls, can be used for all experiments regardless of whether there were replicates. It states that variance increases rapidly from the value observed for the controls as larger sample masses drive the vapor concentration down and the signal area ratio up.

The relationship of Equation VI.25 was also used to fit the linear, Langmuir and B.E.T. isotherms to the sorption on dry adsorption alumina. An example of these results is shown for toluene in Figure VI.10.

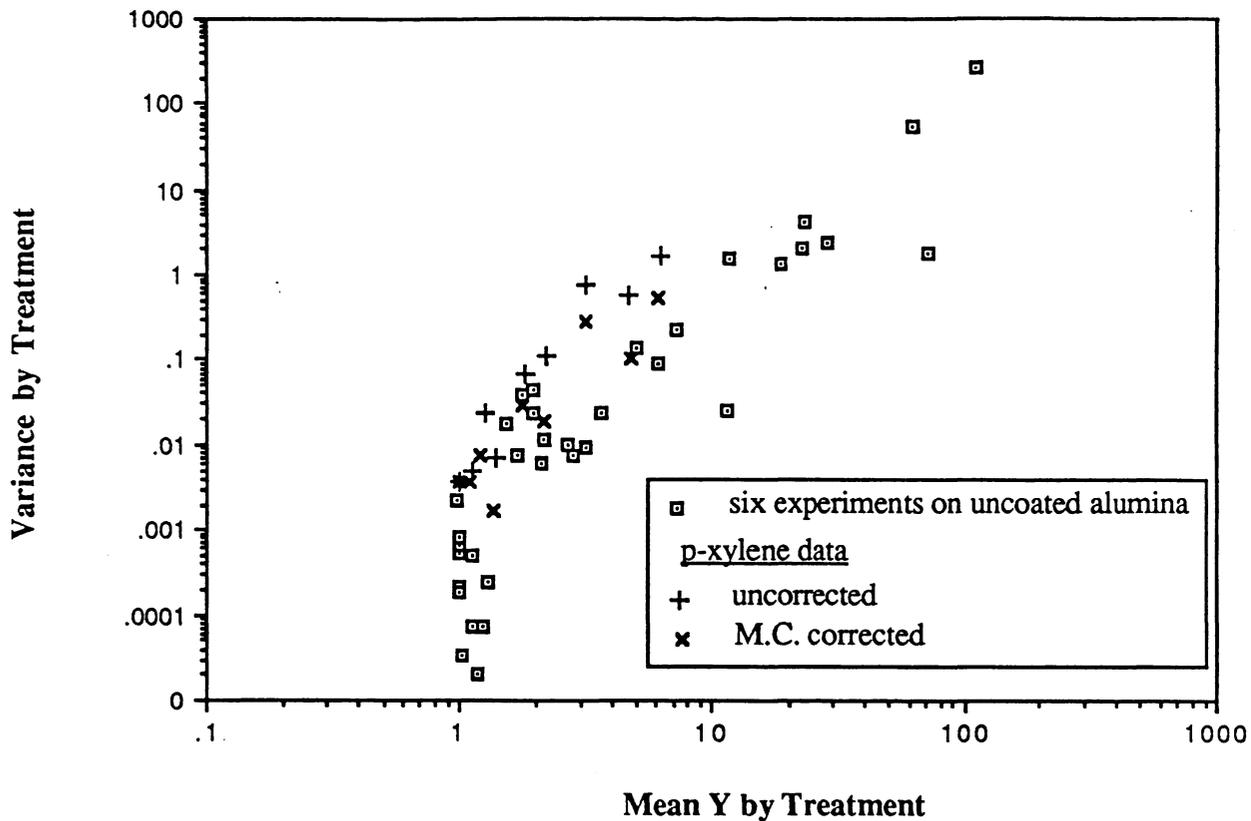


Figure VI.9. Data Used to Fit Equation VI.25, the Variance Model Used for Weighting Laboratory Data.

[Shown are the variances observed for each treatment in six experiments on dry alumina. The dry alumina data is compared to data for p-xylene sorption on coated alumina at 20.4 percent moisture content. The p-xylene data lies on top of the combined data for the dry alumina. After a moisture content correction, the p-xylene data are compatible with the other data.]

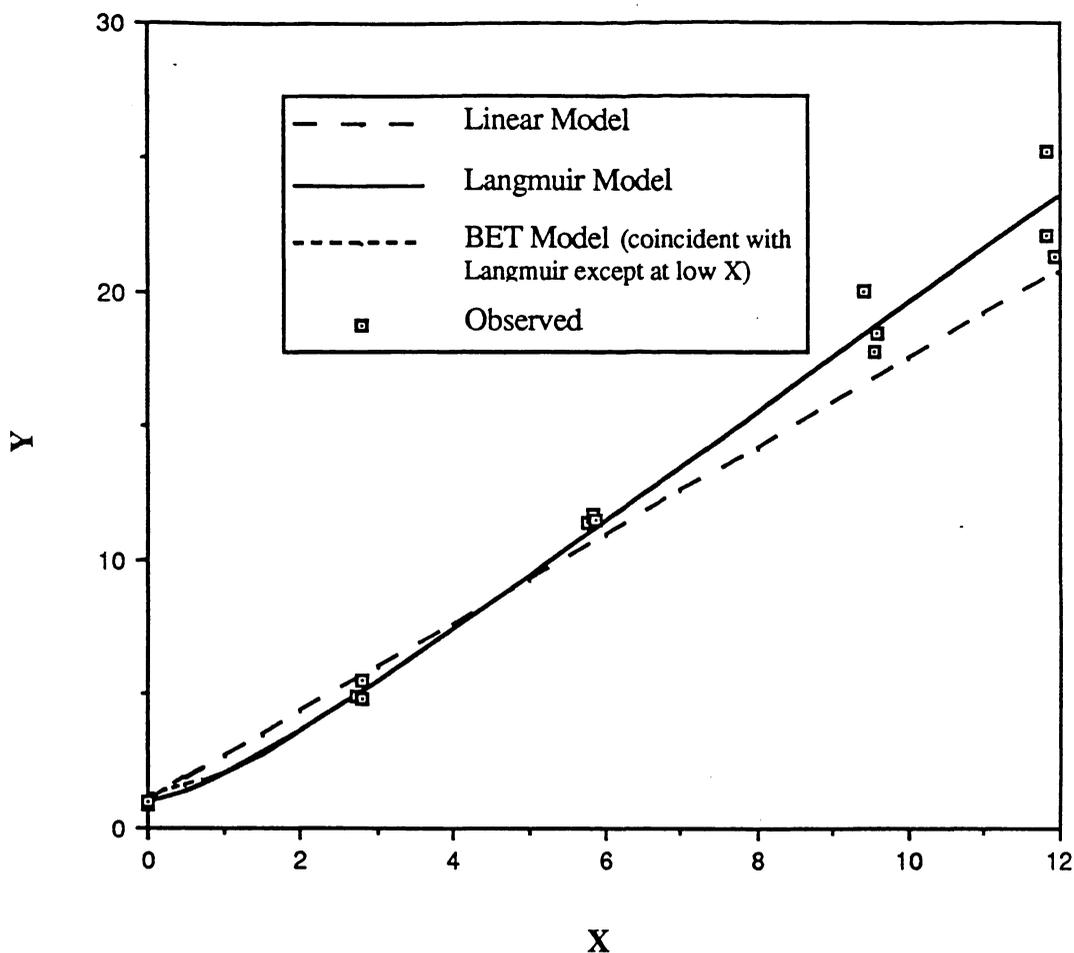


Figure VI.10. Typical data Showing the fit of the Langmuir and B.E.T. Models for Adsorption of Toluene on dry Alumina.

## 2. Vapor Sorption onto Dry Alumina: Model Comparison Study

Sorption isotherms measured by the headspace technique for chloroform, 1,1-dichloroethylene, methylene chloride, perchloroethylene, 1,1,1-trichloroethane, and toluene were fitted to the three models (linear, Equation VI.13; Langmuir, Equation VI.17; and B.E.T., Equation VI.19) as described under theoretical considerations. Values obtained for the fitted parameters and the correlation coefficients are provided (Table VI.4). This work focussed on the statistical analysis of nonlinear isotherms for adsorption concentration data using the Langmuir and B.E.T. models as examples. Because both these nonlinear models both have the same number of adjustable parameters, the weighted residual sums of squares provided sufficient basis for comparison.

TABLE VI.4 FITTED PARAMETERS FOR THE LINEAR, LANGMUIR, AND B.E.T. ISOTHERMS ON DRY ADSORPTION ALUMINA

<u>Vapor</u>	<u>Linear</u>		<u>Langmuir</u>			<u>B.E.T.</u>		
	$K'_d$	$r^2$	$c$	$\Gamma_m$ (mg/g)	$r^2$	$c$	$\Gamma_m$ (mg/g)	$r^2$
CF	2925	0.874	276.4	12.06	0.98	282.0	11.86	0.99
1,1-DCE	120.1	0.955	137.0	1.90	0.99	142.7	1.84	0.99
DCM	208.5	0.885	387.7	4.78	0.94	385.1	4.77	0.94
PCE	1158	0.871	149.9	19.74	0.97	156.5	19.13	0.98
1,1,1-TCA	711.3	0.935	140.1	9.93	0.98	144.4	9.67	0.98
toluene	11670	0.978	122.8	118.5	1.00	125.2	116.2	1.00
TCE	1040	0.998	57.7	19.73	1.00	60.46	18.86	1.00

Equivalent fits were obtained with the Langmuir and B.E.T. equations for the range of vapor concentrations investigated. Figure VI.10 illustrates typical results. Similar values of the parameters  $\Gamma_m$  and  $c$  were predicted by the Langmuir and B.E.T. models. Neither the Langmuir equation nor the B.E.T. equation proved better at fitting the sorption isotherms for dry adsorption alumina for the six compounds investigated. Likely this is due to the fact that the influence of multilayer sorption was negligible at the vapor concentrations used.

Although the B.E.T. model, despite its increased complexity, did not outperform the Langmuir, based on  $r^2$  the two nonlinear models fitted better than the linear case (Table VI.4). The parameter  $\Gamma_m$  was lower than predicted for TCE in Section IV. The value of  $c$  was correspondingly higher than that given in Section IV for TCE. The difference in the values could be explained by the existence of a small number of high affinity sites in the low  $P/P_0$  portion of the isotherm which is the region accessible to the bottle experiments described here. Based on the surface area of the alumina and assuming an approximate molecular surface area of  $30\text{\AA}$ , the value of  $\Gamma_m$  should be about  $100\text{ mg/g}$ . Most values for  $\Gamma_m$  in Table VI.4 are about 20 percent of this maximum value.

Unrealistic estimates were obtained of the parameters  $c$  and  $\Gamma_m$  when nonlinear regression or linearization by Equation VI.3 was applied. Golden Section is a more robust nonlinear regression technique because small perturbations in the data are less likely to misdirect the numerical minimization algorithm. Since the effects of each parameter on the RSS can be separated more effectively by Golden Section, many of the disadvantages of nonlinear regression are eliminated (eg. colinearity in the coefficient matrix used to estimate the parameters). When adsorption is measured indirectly through concentration differences,

the method proposed here will provide better estimates of the fitting parameters than nonlinear regression.

## F. SUMMARY AND CONCLUSIONS

Fitting adsorption data to adsorption isotherms allows hypotheses about the nature of the adsorption process to be tested. When the extent of sorption is inferred from the concentration difference before and after the sorption process, isotherm fitting approaches must reflect the pattern of error expected in this type of data. Errors may be large, especially when the concentration difference observed, as well as the mass of sorbent, are small.

The Golden Section approach provides a method of using weighted least squares to incorporate the pattern of errors commonly observed for adsorption measurements made by concentration difference. Instead of a difference as the response variable (Equation VI.11), a ratio  $Y$  (Equation VI.12) is used to compare the concentration after adsorption to a reference concentration, representing what was present initially before adsorption. Nonlinear equations for the Langmuir and B.E.T. isotherms were fitted to the data using a weighted nonlinear numerical optimization procedure. Weights, based on an exponential model for the variance, were derived from a separate experiment. The same values for the two parameters,  $\Gamma_m$  and  $c$ , were predicted for both the Langmuir and B.E.T. isotherms. This result is expected for data in the low vapor concentration range which was examined here.

Current nonlinear regression practices for fitting the parameters of complex adsorption isotherms are questioned for the data in the low concentration range of the isotherm. When data in the linear, or near linear range is measured by concentration level difference, the linear isotherm (Equation VI.13) was better suited for regression analysis because concentration became the regressor. Likewise this approach allows the variable known with the least error, sorbent mass, to function as the predictor. Overall, slightly better parameter estimates were obtained with Equation VI.13 than Equation VI.9.

Since nonlinear regression was hampered by correlation between the parameters, the Golden Section method was superior for parametrizing isotherms to experimental data. Golden Section did the best job at parametrizing a synthetic data set compared to linearized Equation VI.3 and nonlinear regression (Equation VI.1). Weighting the data allowed the fitted curves and parameters obtained to reflect what is known about the level of certainty at each data point. Weighting improved the consistency of predictions by each method by correcting for the effect of transformations on the  $Y$  values. Weighting was unable,

however, to control for the errors which occur in the predictor variables X, which explains why linearizations of the Langmuir equation (Equation VI.3) and nonlinear regression on Equation VI.1 did not perform as well as the Golden Section method. The agreement between the Langmuir and the B.E.T. isotherm fits for these data also supports the validity of the Golden Section regression method.

## G. FINAL COMMENTS

Two choices confront the researcher preparing to analyze adsorption data: 1) whether to use a linear or nonlinear (eg, Langmuir, B.E.T., etc.) isotherm; and 2) once a particular model is chosen, deciding which form of the model equation to use. Both of these issues have been addressed by Kinniburgh (1986) for the case of adsorption data collected at a fixed sorbate concentration. Kinniburgh, however, discussed cases where the amount adsorbed per gram is measured directly rather than calculated from the observed adsorbate concentration difference, as is often the case.

For the first issue, ie. that of whether to fit a one or two parameter model, Kinniburgh (1986) suggested the following equation to test whether or not to remove a parameter from the model. This equation applies to comparisons between concentric models, ie. where the model with fewer parameters is a limiting case of the model with more parameters.

$$F(p_2 - p_1, m - p_2) = \frac{(RSS_1 - RSS_2) / (p_2 - p_1)}{RSS_2 / (m - p_2)} \quad \text{VI.46}$$

where  $F(p_2 - p_1, m - p_2)$  is the F statistic with  $(p_2 - p_1, m - p_2)$  degrees of freedom,  
 $p_1$  = the number of parameters in the limiting model, ie.  $p_1 = 1$  for the linear isotherm,  
 $p_2$  = the number of parameters in the more complex model, ie.  $p_2 = 2$  for the Langmuir model,  
 $m$  = the number of data points or observations taken,  
 $RSS_1$  = the residual sum of squares of the limiting model, and  
 $RSS_2$  = the residual sum of squares of the more complex model.

For the second issue (which is the focus of the discussion in this section), Kinniburgh (1986) showed that all forms of the Langmuir isotherm equation resulted in essentially the same values for the parameter estimates, when weighted linear regression was used (provided the weights were assigned correctly). Kinniburgh (1986) gives proper weights for a variety of different isotherms, including the Langmuir and B.E.T. isotherms. Many

adsorption isotherms, however, such as the Redlich-Peterson isotherm, are more complex than the Langmuir or B.E.T. isotherms. For the complex isotherms, involving more than two fitting parameters, Kinniburgh advocated using nonlinear regression.

The weights given by Kinniburgh, unfortunately, are not always sufficient for analysis of the adsorption data where the amount of adsorption is inferred from adsorbate concentration difference. This was the case in the headspace analysis technique in this study. Data collected by the headspace analysis technique, or other techniques where adsorbate concentration difference is measured, produce errors associated with both the predictor variable (adsorbate concentration) and the regressor variable (amount of adsorption). The weighting approach advocated by Kinniburgh accounts only for errors in the response variable. When adsorption is measured by concentration difference, it is therefore necessary to express the equations for the Langmuir and B.E.T. isotherms in a form where the standard weighting equations used by Kinniburgh are valid. Since the variables were not separated in the resulting equations, a Golden Section search routine was employed for their solution. Analysis of synthetic data (Table VI.1) demonstrated that the weights suggested by Kinniburgh performed poorly for Equation VI.1 and Equation VI.4. Therefore, it is suggested that the Langmuir equation be linearized by Equation VI.3 or that the golden section approach be used on Equation VI.17. Additionally, for fitting linear data, it is suggested that Equation VI.13 be used rather than Equation VI.9. Figures VI.1-VI.8 compare parameters estimates by several of the best approaches. The data analysis methods presented here are used to aid in the comparison of the sorption of different vapors both as individual compounds and as mixtures in Sections VII and VIII.