

Linear Multiple Regression Model of High Performance Liquid Chromatography

STANISLAVA LABÁTOVÁ

Institute of Informatics

Department of discrete processes modeling and control

Slovak Academy of Sciences

Dúbravská 9, 845 07 Bratislava

SLOVAKIA

utrrlabt@savba.sk, <http://www.sav.sk>

Abstract: - The multiple regression analysis is proposed as method for presentation results of high performance liquid chromatography (HPLC), which simultaneously depends on several parameters. The coefficients of linear multiple regression model are assessed by the least square method. In this way dependent variable – retention volume – is expressed by more than one explanatory variable (i.e. by the molar mass, concentrations of eluent and temperature). The approach was to depict dependences of retention volumes on molar mass of polystyrene probes in tetrahydrofuran (THF) and in mixed eluents THF plus dimethylformamide eluents on the silica gel alkyl bonded column packing. The dependences retention volumes on molar mass, concentrations of eluent and are shown.

Key-Words: -Multiple regression analysis, model specification, high performance liquid chromatography, calibration curves

1 Introduction

High performance liquid chromatography (HPLC) belongs to the most important analytical tools of modern chemistry. HPLC separates complex mixtures of both low and high molecular species to allow conclusions on sample composition in terms of chemical structure and architecture of components and also in terms of their molar mass averages and distributions. HPLC is based on differences in retention of separated substances within appropriate stationary phase. The analyzed phase is introduced into device (usually a column) filled with stationary phase (eluent). The heart of all HPLC systems is the column packing. The volume of mobile phase needed to eluate the particular substance from the column is measured. This is the retention or elution volume (V_R) and reflects the extent of retention of analyze substance within the column packing. Concentration of analytes leaving the HPLC column is monitored by means of detectors, which measure light absorption

of the column effluent. The dependence of the detector response on eluent volume is shown on a *chromatogram*.

Many real polymers consist of different chemical macromolecular components. These are mixed intentionally to obtain the expected system characteristics or they arise during the use of polymers (i.e. light degradation). A de-formulation intentionally prepared complex polymers mixtures to specify the role of individual elements of system and to try to improve their characteristics in next. Analysis and molecular characterization of system element obtained during construction or degradation poly-reactions allows better understanding and verification. The aim of degradation reaction is exact knowledge of process mechanism that allows in the end to suggest the better systems.

Direct analysis of chemical structure of complex polymers and especially of minor (<1%) components in multi-component polymer blends is

not possible by the classical analytical methods. New unconventional methods of separation of complex multi-component polymer system with emphasis on minor (<1%) components are needed. Selected methods of high performance liquid chromatography of polymers, (polymer HPLC) belong to the most potential candidates. In polymer HPLC, the column packings are usually formed with beds of porous particles. Dependences of V_R data on molecular characteristics of polymer (s) under study are evaluated. The most common polymer HPLC method is size exclusion chromatography, SEC. SEC is based on the dependence of elution rate of macromolecules on their size. Large macromolecules permeate only few pores of column packing and their progression along the column is fast. At the same time, smaller species permeate many pores and their elution rate is low. As a result, macromolecules are separated according to their size. One speaks about exclusion retention mechanism. However, it is often necessary to separate macromolecules also according to their chemical composition or physical architecture. The problem is that at least two molecular characteristics simultaneously affect the retention volume of a complex polymer, namely size (molar mass) and chemical composition or architecture of macromolecules. It is necessary to discriminate these two effects in order to independently determine two or even three molecular characteristics of a polymer sample. To do so, two or even several retention mechanisms are combined within one single column or within a series of different columns. The most potential retention mechanisms alternative to size exclusion are based on differences in the adsorption of macromolecules on surface of column packing or in the enthalpic partition of macromolecules between the volumes of stationary phase and mobile phase. Correspondingly, one speaks about adsorption and enthalpic partition retention mechanism of separation. The pore exclusion is, however always present as an additional retention mechanism. Consequently, one deals with the coupled retention mechanisms adsorption plus exclusion or enthalpic partition plus exclusion. The aim of coupling of retention mechanisms in polymer HPLC is to either enhance or to suppress separation of macromolecules according to one molecular characteristic so that the effect of other characteristic(s) can be neglected. The results obtained with coupled methods of polymer HPLC are complex and it is difficult to discriminate effect

of particular retention mechanisms on the resulting retention volumes.

Experimental results are mathematically processed in order to obtain improved description and generalized elucidation of retention mechanism. The aim is to describe a system with optimal number of parameters, which would allow generalized conclusions and predicting of the behavior of polymer systems.

Novel approaches to liquid chromatography separation of synthetic polymers were elaborated, which combine two or several separation mechanisms. These approaches are called coupled method of polymers high performance liquid chromatography. The aim is to either enhance or suppress separation of macromolecules according to one molecular characteristic so that the effect of other characteristic(s) can be neglected. Unconventional liquid chromatography methods of separation of complex multi-component polymer systems [3] are mathematically processed in order to obtain model for HPLC process and generalized elucidation of the retention mechanisms.

The remainder of the paper is organized as follows: Section 2 deals with calibration curves. Section 3 explains model specification using multi regression analysis. Section 4 contains examples of linear multi regression model and in Section 5 are done conclusions.

2 Optimal Design of Calibration Curves for HPLC by Polynomial Regression

One important step in the data treatment of HPLC experiment is modeling of calibration curves (CC). The dependence retention volume on molar mass is called calibration curves. The CC –s are generally not linear.

From mathematical point of view for solution above problem as a dependence retention volume on molar mass (CC) we can write:

$$Y_i = f(x_i) + \varepsilon_i; \quad i=1,2,\dots,n, \quad (1)$$

where

x_i is value of independent variable of HPLC experiment, i.e. molar mass

Y_i is measured response of HPLC experiment – retention volume

f is unknown function

ε_i is random error

n is a number of measurements

We would be willing to assume that function f is sufficiently smooth over the range of interest, so we represent the function f adequately by a polynomial. If we assume that function

$$f(x_i) = P_m(x_i) =$$

$$a_0 + a_1x_i + a_2x_i^2 + \dots + a_mx_i^m, \quad i = 1, 2, \dots, n, \quad (2)$$

is a polynomial of degree m , where a_i are unknown, we will call the setup model P_m . Let Y, a, ε denote column vectors, (transpose line vectors $Y' = (Y_1, Y_2, \dots, Y_n)$, $a' = (a_0, a_1, \dots, a_m)$, $\varepsilon' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$) and X denote a matrix of type $(n)/(m+1)$ given by:

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \Lambda & x_1^m \\ 1 & x_2 & x_2^2 & \Lambda & x_2^m \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \Lambda & x_n^m \end{bmatrix} \quad (3)$$

then the model P_m can be written as :

$$Y = Xa + \varepsilon, \quad (4)$$

where x_i 's take at least $m+1$ distinct values and $n \geq m+1$. Then unique least squares estimator of a is given by

$$a^* = (X'X)^{-1}XY = (a_0^*, a_1^*, \dots, a_m^*). \quad (5)$$

Question is, which criterion can be considered as compromises between the incompatible goals of inference about regression function under an assumed model and checking model adequacy.

For evaluation estimated regression function (in our case polynomial m order) is possible to use several ways [4]:

1. Coefficient of multiple determination (R^2)
2. Test based on analysis variance
3. The root mean square error
4. An analysis of residuals.

Criterion that we used is application root mean square error (RMSE) method given by

$$E = \sqrt{\frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}}. \quad (6)$$

Using some level of importance it is possible to define compromise optimal criterion, which detects information about exactness of parameters and share of model, respectively. Different polynomial models have been evaluated and compared, not only for model fit but also for their predictive properties.

Table 1
Kolóna Kromasil C-1, 100A, 150x7.8 mm

log M	Mw	100 % THF	20 % THF	18 % THF	17.5 % THF	17 % THF	100 % DMF
2.82	666	3,94	4.62	4.71	4.79	4.73	4.82
3.41	2600	3,5	4.35	4.38	4.48	4.52	4.67
3.95	9000	3,1	3.96	3.95	4.03	3.97	4.37
4.24	17500	2,79	3.36	3.44	3.42	3.45	3.88
4.57	37000	2,71	2.91	2.97	2.96	2.98	3.37
5	100000	2,67	2.7	2.69	2.73	2.74	2.71
5.37	233000	2,68	2.69	2.7	2.7	2.73	2.7
5.7	498000	2,67	2.65	2.67	2.67	2.73	2.66

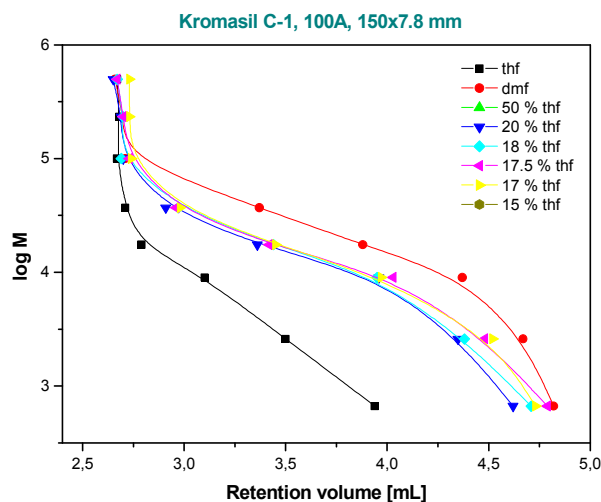


Fig.1 Dependences molar mass on retention volume

In HPLC, magnitude of parameter which represents molar mass, is too large so from practical and historical arguments in next there is considered with their logarithm.

For creating model describing calibration dependences in HPLC in first we suppose that the empirical formula has the form:

$$M = Ae^{P_m(x)}, \quad (7)$$

and

$$\log M = \log A + P_m(x). \quad (8)$$

An empirical function we wanted to have minimal (for established polynomial coefficients by using least square method) and have the form:

$$\Phi(a_0, a_1, \dots, a_m, \log m) = \sum_{i=1}^n [a_0 + a_1 x_i + \dots + a_m x_i^m - \log M]^2 \rightarrow \min \quad (9)$$

so

$$\frac{1}{2} \frac{\partial \Phi(a_0, a_1, \dots, a_m, \log M_i)}{\partial a_i} = \quad (10)$$

$$\sum_{i=1}^m (a_0 + a_i x_i^i - \log M_i) = 0.$$

Solving this system of linear equations we obtain polynomial coefficients a_0, a_1, \dots, a_n .

Dependence molar mass on retention volume at concentration of eluent which is equal 0% THF, for experimental data from Tab.1, estimated by polynomial of 2nd order, is given by:

$$Y = 0.22x^2 - 2.33x + 8.81 + \varepsilon_2. \quad (11)$$

The root mean square error: $\varepsilon_2 = 0.05585$.

See Fig.2

Dependence molar mass on retention volume where the concentration of eluent is equal 0% of THF, for experimental data from Tab.1, estimated by polynomial of 3rd order is given by:

$$Y = 0.04x^3 - 0.3x^2 - 0.17x + 5.93 + \varepsilon_3. \quad (12)$$

RMS error: $\varepsilon_3 = 0.0506$

See Fig. 2

For the polynomial model of 5-th order for experimental data from Tab.1 we obtain next coefficients:

$$Y = -0.04x^5 + 0.76x^4 - 5.6x^3 + 19.7x^2 - 34.03x + 27.6 + \varepsilon_5, \quad (13)$$

RMSE: $\varepsilon_5 = 0.02841$.

Graf of estimated polynomial functions (2nd, 3rd, 5th orders) for experimental data from Tab.1 are shown on Fig.3.

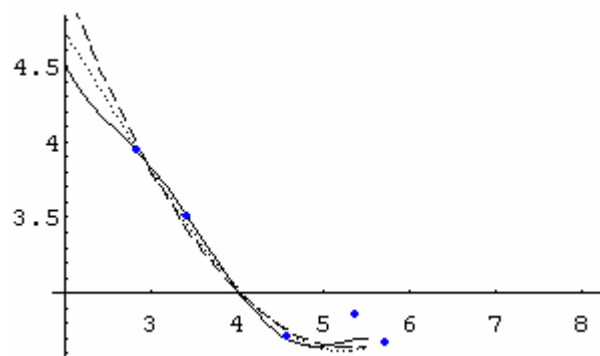


Fig.2 Calibration curves for experimental data from Table 1., where

..... represents polynomial 2nd order,
--- represents polynomial 3rd order,
_____ represents polynomial 5th order

Evidently, for every value of eluent concentration we obtain different calibration curve, which is very immense and for presentation results of HPLC experiment non available.

System for evaluation of calibration curves in liquid chromatography is constructed as Multi – Agent System (MAS). System consists from decision agent and agent for modeling and processing different cases of CC. In this work the agent for processing CC is P_3 most reflected, as this agent is very suitable for our aim. The shape of cubic parabola depends on value $D = 3a_2a_0 - a_1^2$. If $D < 0$ then polynomial P_3 has one maximum and one minimum, which characteristic is used for analysis for example for obtaining information about tendencies in behavior of macromolecules in HPLC.

For presentation results HPLC is important to observe not only CC, that represents only dependencies two variables (molar mass and retention volume), but also influence of next variables that evidently to influence on process HPLC. For observing multiple dependencies that have influence on process HPLC a multiple regression is a suitable tool.

3 Model specifications using multiple regression

Multiple regression (MR) analysis is a statistical tool for understanding the relationship between two or more variables in a system. In other way, MR is a regression giving conditional expectation value of given variable in terms of two or more variables. Generally, MR involves a variable to be explained – called the dependent variable – and additional explanatory variables that are thought to produce or be associated with changes in the dependent variable. The important variables that systematically might influence the dependent variable, and for which data can be obtained (measured), typically should be included in the model. All remaining influences, which should be small individually, but can be substantial in the aggregate, are included in an additional random error term.

MR is a procedure that separates the systematic effects (associated with the error term) from the random effects and also offers a method of assessing the success of the process.

MR analysis is sometimes well suited for analysis of data about competing theories. There are several possible explanations for the relationship among a number of explanatory variables to assess the statistical data pertinent to these theories.

MR regression also may be useful in:

1. Determining whether or not a particular effect is present
2. Measuring the magnitude of particular effect
3. Forecasting what a particular effect would be, but for an intervening event.

Moreover, in interpreting the results of MR analysis, it is important to distinguish between correlation and causality. Two variables are correlated when the events associated with variables occur more frequently together than one would expect by chance. Correlation between two variables does not imply that one event causes the second to occur. Therefore, in making causal inferences, it is important to avoid spurious correlation.

MR allows the expert to choose among alternative theories or hypothesis and assists the expert in sorting out correlations between variables

that are plainly spurious from those that reflect valid relationships.

Model specification involved several steps, each of which is fundamental to the success of the research effort. Models are often characterized in terms of the parameters – numerical characteristics of the model.

MR uses a sample to obtain estimate of the values of parameters of the model – an estimate associated with a particular explanatory variable is a regression coefficient.

Multiple regression analysis goes beyond the calculation of correlations; it is a method in which a regression line is used to relate the average of the variable – the dependent variable – to the values of other explanatory variables. As a result, regression analysis can be used to predict the values of one variable using the values of others.

Steps for model specification:

1. Choosing the dependent variable – the variable to be explained should be appropriate for analyzing the question in the issue.
2. Choosing the explanatory variable that is relevant to the issues in the case. The explanatory variable that allows the evaluation of alternative hypotheses must be chosen appropriately.
3. Choosing the additional explanatory variables – an attempt should be made to identify the additional known or hypothesized explanatory variables, some of which are measurable and may support alternative substantive hypotheses that can be accounted for by the regression analysis.
4. Choosing the functional form of the MR model – the expert must also choose the proper form of the regression model, most frequently linear form, or rare nonlinear form.

Regression results can be interpreted in purely statistical terms, through the use of significance tests, or they can be interpreted in a more practical, non-statistical manner [2]. In our case, we use the latter.

4 Linear multi regression model of HPLC

In general, linear multi regression model takes the following form:

$$Y = k_0 + k_1X_1 + k_2X_2 + \dots + k_nX_n + \varepsilon, \tag{14}$$

where Y represents the dependent variable (i.e. the variable to be explained),

X_1, X_2, \dots, X_n represent the explanatory variables and

ε – error, which represents the collective unobservable influence of any omitted variables. Coefficients k_0, k_1, \dots represent multiple regression coefficients which are estimated by fitting the equation to the data by some appropriate method.

Using the unconventional HPLC method we obtain experimental data set

$$\{V_{Ri}, \log M_i, THF_i\} \quad i = 1, 2, \dots, n, \tag{15}$$

where V_R is retention volume
 $\log M$ is log molar mass
 THF is concentration of eluent

For creating a model for HPLC in first we suppose that empirical formula has the form:

$$V_R = k_1 * \log M + k_2 * THF + k_0. \tag{16}$$

For determination of coefficients k_0, k_1, k_2 from (3) we minimize function Φ given by discrete set of points (2) (least square method):

$$\Phi(k_0, k_1, k_2) = \sum_i [k_1 \log M_i + k_2 THF_i + k_0 - V_{Ri}]^2 \rightarrow \min. \tag{17}$$

$i = 1, 2, \dots, n$

So we obtain system of linear equations:

$$\frac{\partial \Phi(k_0, k_1, k_2)}{\partial k_0} = 0; \quad \frac{\partial \Phi(k_0, k_1, k_2)}{\partial k_1} = 0;$$

$$\frac{\partial \Phi(k_0, k_1, k_2)}{\partial k_2} = 0. \tag{18}$$

Solution of this system of linear equations determines coefficients k_0, k_1, k_2 .

The least squares regression provides not only parameter estimates that indicate the direction and

magnitude of the effect of the change in explanatory variable on the dependent variable but also an estimate of the reliability of the parameter estimates and measure of the overall goodness-of-fit of a regression model.

Experimental data for polystyrene in neat *THF* and mixed eluent *THF + DMF* are shown in Table 2:

Kromasil C-18, 100A, 2x150x7.8 mm

log M	Mw	100 % THF	20 % THF	19 % THF	18 % THF	17 % THF	100 % DMF
2.82	666	7.2	10.03	10.21	10.15	11.00	13.33
3.41	2600	6.2	10.11	10.62	10.66	11.48	
3.95	9000	5.66	10.03	10.79	11.07	-	
4.24	17500	5.32	9.55	10.57	11.35	12.55	
4.57	37000	5.25	8.43	10.03	11.25	12.75	
5	100000	5.18	5.75	7.17	9.37	12.06	
5.37	233000	5.16	5.22	5.81	8.55	11.41	
5.7	498000	5.17	5.07	5.12	5.67	-	

Table 2
experimental data

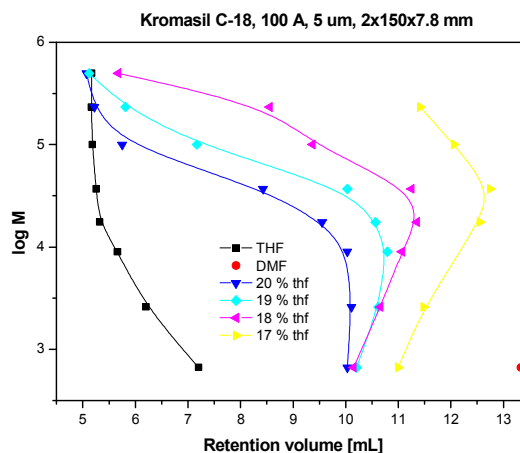


Fig.3 Dependences of retention volume on molar masses – calibration curves for different concentration of eluent

The retention volume V_R is the subject of our interest of explanation, and thus it has been chosen as dependent variable. Explanatory variables are choosing next variables: the percentage of

concentration of eluent and molar masses. We intend to estimate the dependent variable V_R in the form where exist two explanatory variables ($\log M$, THF):

$$V_R = k_0 + k_1 \log M + k_2 THF + \varepsilon. \quad (19)$$

This suggested equation is now fitted by appropriate method for estimate regression coefficients. Using the least square method for estimate regression coefficient we obtain the results as shown in Table 2:

M	T	V	NV
2.820	1.000	7.200	7.640
3.410	1.000	6.200	6.863
3.950	1.000	5.660	6.152
4.240	1.000	5.320	5.770
4.570	1.000	5.250	5.335
5.000	1.000	5.180	4.769
5.370	1.000	5.160	4.282
5.700	1.000	5.170	3.847
2.820	0.200	10.030	11.430
3.410	0.200	10.110	10.653
3.950	0.200	10.030	9.942
4.240	0.200	9.550	9.560
4.570	0.200	8.430	9.125
5.000	0.200	5.750	8.559
5.370	0.200	5.220	8.071
5.700	0.200	5.070	7.637
2.820	0.190	10.210	11.477
3.410	0.190	10.620	10.700
3.950	0.190	10.790	9.989
4.240	0.190	10.570	9.607
4.570	0.190	10.030	9.172
5.000	0.190	7.170	8.606
5.370	0.190	5.810	8.119
5.700	0.190	5.120	7.684
2.820	0.180	10.150	11.525
3.410	0.180	10.660	10.748
3.950	0.180	11.070	10.036
4.240	0.180	11.350	9.654
4.570	0.180	11.250	9.220
5.000	0.180	9.370	8.653
5.370	0.180	8.550	8.166
5.700	0.180	5.670	7.731
2.820	0.170	11.000	11.572
3.410	0.170	11.480	10.795
4.240	0.170	12.550	9.702
4.570	0.170	12.750	9.267
5.000	0.170	12.060	8.701
5.370	0.170	11.410	8.213

$e = 1.658$
 $k_1 = -1.3170526$
 $k_2 = -4.7372490$
 $k_0 = 16.0913580$

Table 3 Results of multiple regression estimation for experimental data from Table 1

where

NV - is an estimate retention volume
 k_0, k_1, k_2 - are estimate regression coefficients,
 e - is additional random error,
 M - represents a log of molar mass,
 THF - is eluent concentration
 V_R - is retention volume.

Using the results from Table 1 we can express the V_R as follows:

$$V_R = -1.31705 M - 4.7372 THF + 16.0913 + \varepsilon, \quad (20)$$

where $\varepsilon = 1.658$ (additional random error) is connected with variables which are not included in empirical form (19) and which individually have small influence on dependent variable (retention volume), but can be substantial in aggregation.

The results of linear multiple-regression for experimental data from Tab.1 expressed by (20) are shown in Fig. 4.

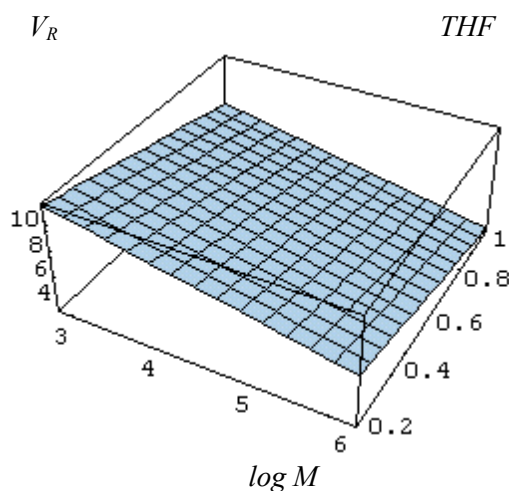


Fig. 4 Dependences of retention volume on molar masses and concentrations eluent THF

The random errors of regression coefficients tell us how much the parameter estimate is likely vary from sample to sample.

Kromasil C-18, 100A; 150x7.8 mm

(temperature influence)

log M	Mw	100 % THF	50 °C	40 °C	35 °C	31 °C	30 °C
2.82	666	3,43	4,48	5,1	4,93	4,89	5,54
3.41	2600	3,01	4,6	5,09	5,17	5,15	5,94
3.95	10100	2,68	4,31	5,11	5,43	5,59	6,42
4.24	17500	2,63	4,15	5,05	5,49	5,82	6,87
4.57	37000	2,58	3,55	4,68	5,2	6,26	7,72
5	100000	2,58	2,69	3,33	4,4	7,24	
5.37	233000	2,56	2,58	2,79	3,7		
5.7	498000	2,52	2,55	2,71	2,87		
6.08	1200000	2,52	2,53	2,69	2,66		

Table 4 HPLC experimental data for influence temperature on dependency molar mass on retention volume

Linear multiple regression model for explaining dependence molar mass and temperature on retention volume can to use empirical form:

$$V_R = k_0 + k_1 \log M + k_2 T + \varepsilon \quad (21)$$

where

- $\log M$ represents molar mass
- T represents temperature
- V_R represents retention volume
- k_0, k_1, k_2 represents regression coefficients
- ε represents additional random error

Using experimental data from Tab.4 for temperature influence, HPLC process we can express by equation:

$$V_R = -0.654 M - 0.106T + 11.551 + \varepsilon, \quad (22)$$

$$\varepsilon = 0.854$$

The additional random error ε is connected with events that exist in the process HPLC, but they are not taken in empirical form (as they magnitudes are not relevant, or for which does not exist experimental data).

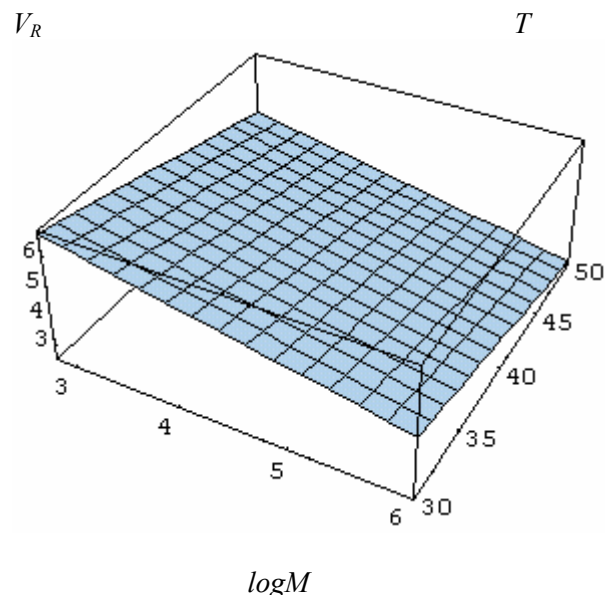


Fig.5 Linear multiple regression model for dependences temperature and molar mass on retention volume

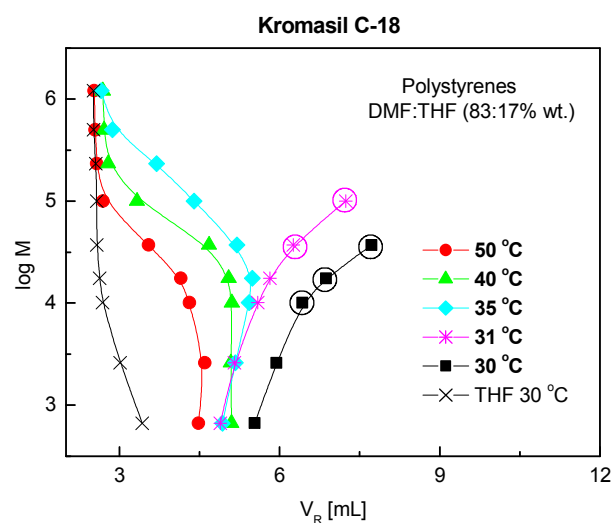


Fig. 6 Dependency retention volume on molar mass for different temperatures (calibration curves for different temperatures)

The more complicated model, where we intent to suppose that linear model contain three explanatory variables (molar mass, concentration of eluent, temperature) can have the form:

$$V_R = k_0 + k_1 \log M + k_2 THF + k_3 T + \varepsilon \quad (23)$$

Again, the error ε (additional random error) is connected with other events that are not expressed in (22). In future we intend to include further additional explaining variables that express V_R and that have the influence (we suppose that small) on interpretation of the experiment. In [3] V_R is expressed only as polynomial dependence on $\log M$. It is evident that the process of HPLC is the function of several other additional variables. On the other hand, the dependency of retention value on molar masses is in general nonlinear, so we intend to use a more precise as previous explained linear model a nonlinear MR model for HPLC in our future work.

5 Conclusion

High Performance Liquid Chromatography separates complex mixtures of molecular species to allow conclusions in terms of their chemical structure, architecture and molar mass distributions. As the direct analysis of chemical structure of complex multi-components polymers and especially of minor (<1%) components is not possible, new unconventional methods of separation are needed. Selected methods of HPLC belong to this class of methods. Unconventional liquid chromatography methods of separation of complex multi-component polymer were mathematically processed by experimental data for elucidation of the retention mechanisms.

Multiple regression analysis is a tool for understanding the relationship in system to be observed. Multiple regression analysis is well suited to the analysis of data competing alternative theories or hypothesis in which exist several possible explanations for the relationship among a dependent variable and number of explanatory variables. Problem is, on the other hand, how to obtain experimental data to be useful for our assumptions.

For understanding process of HPLC at first there were studied the relation between molar mass and retention volume – which is represented by calibration curves. Calibration curves were expressed by polynomials of different orders for different values of concentrations of eluent. In the past, dependencies retention volume had been studied on one variable only – the molar mass. It was realized by polynomial regression of the 2nd and 3rd or 5th order. By approximations calibration

curve does not be solved problem of HPLC process – calibration curves are only one part of HPLC that explain process of HPLC. Polynomial approximations of calibration curves are diametrical different for individual concentrations, so we cannot detect how they change by influence of concentration of eluent. For our goal – to describe process of HPLC it is insufficient. This question is partly solved by linear multiple regression model.

Using multiple regression to establish regression coefficients for expressing empirical functional form for retention volume in HPLC represents a next step of explanation dependent variable as the dependence on more than one explanatory variables. In HPLC many different events is present that have some influence on this process and that have not been considered in the previous case. In (19) in the empirical functional form explanatory variables are included: molar mass and eluent concentration. It is evident that in process HPLC are present more than two explanatory variables. This reality is connected with error, explained in (20). This error can be decreased when the empirical functional form for expression process HPLC will be expressed with more explanatory variables (molar mass, concentrations of eluent, temperature,...) that may have a small influence individually but in aggregation may be important. Sometimes is a problem to obtain experimental data for next explanatory variables.

Indeed, calibration curves that are basic in HPLC process are not generally linear, so an empirical function for multiple regression will be interesting to suggest in non-linear form. Experiences with polynomial regression approximations of calibrations curves there will be used. Solution of non-linear multiple regression model is not trivial problem.

In the future we will deal with next events and additional variables that may have some importance in problem HPLC by a non-linear multiple regression model.

Acknowledgements: Financial support from the APVV Slovak Grant Agency (project 0592– 07). The author is thankful to Dr. D. Berek for his advice and to Mrs. J. Tarbajovska for her excellent technical assistance.

References

- [1] I. Rustanov, T.Farcas, F.Ahmed, F.Chan, R LO Brutto, H.M.Mc Nair, Y.V. Kazakevich, *J.Chromatogr. A* 2001,913, 49
- [2] Reference manual on Scientific Evidence
<http://air.fjc.gov/public/fjc.nsf/pages/16>
- [3] D.Berek, S. Labatova, *WSEAS Proc. Studies in Simulation and Modeling, Vancouver, 2007*, pp.91 – 96
- [4] D.L.Massart, B.G.M. Vandeginste, L.M.C.Buydens, S.De Jong, P.J.Lewi, J.Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier, Amsterdam, 1997