**Data Pre-processing for Equilibrium Dispersive Model Parameters Determination Modeling Preparative Chromatography Separation of Mannitol, Glucose, Fructose, and Sucrose from the Mixture**

T. Svoboda, S. Henke, S. Gillarová, D. Růžička, M. Sluková

University of Chemistry and Technology Prague, Faculty of Food and Chemical Technology, Department of Carbohydrates and Cereals, Technicka 5, 166 28 Prague 6, Czech Republic;

tel. +420 775 220 977, email: tomas3.svoboda@vscht.cz

Preparative chromatography using ion exchange resin is efficient method for the separation of monosaccharides. Continuous chromatography with a simulated moving bed is a popular technology for industrial production of pure chemicals because of its efficiency in the terms of sorbent utilization and eluent consumption. To find the optimal operating parameters for the SMB chromatography system, a mathematical description of the chromatography process is necessary.[3], [4]

The Equilibrium Dispersive Model (EDM) extends the simple equilibrium model published by Don De Vault in 1943.[1] He described chromatography as a concentration wave propagation in a fixed bed column with assumptions of diluted systems, plug flow, negligible pressure drop, and isothermal operation. Addition of second-order derivative apparent diffusion (dispersive) term is useful to model nonlinear behavior which manifests itself, for example in peak asymmetry. The mass balance of the component *i* in a volume element of the column is

|  |  |  |
| --- | --- | --- |
|  |  | *equation 1* |

Combining with the linear adsorption equilibrium isotherm.[3],[4]

|  |  |  |
| --- | --- | --- |
|  |  | *equation 2* |

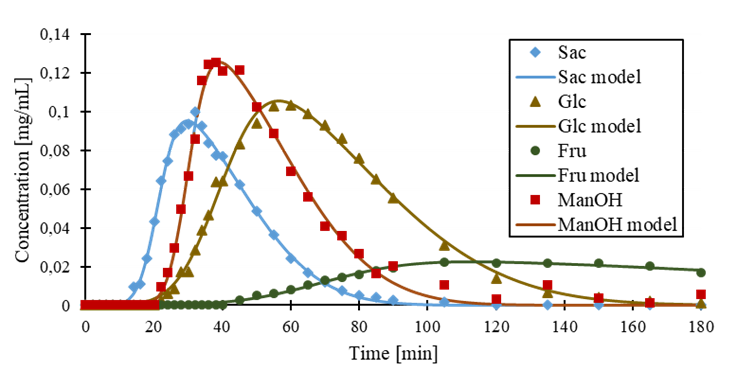
This study presents the original measured data from discontinuous preparative chromatography pulse feed experiments and introduces the methods of pre-processing of the data and way to utilise them for determination of parameters of the EDM with non-competitive linear isotherm.

The mixture of mannitol, glucose, fructose, and sucrose was separated during the experiments. The experimental setup consisted of the column XK16/40 (GE, USA) with a commercially accessible Dowex 550A anion exchange resin (Dow, USA) with a layer height of 23.5 cm and a diameter of 1.6 cm. The discontinuous station also consisted of a piston dosing pump and a three-way electromagnetically controlled valve used to inject the separated mixture. The fraction samples from these separation experiments were analysed on HPAEC (High-Performance Anion Exchange Chromatography) Dionex equipment. It is an ion-exchange liquid chromatograph with pulsed amperometry detection. The concentrations of the components in the individual fractions were determined by integration of the peaks and comparison with the standards.

Experimental datapoints are loaded with errors arising from sample pipetting, integration of analytical peaks, etc. Therefore, pre-processing algorithms have been proposed. The first step is the replacement of the original measured datapoints with the fitted asymmetric Gauss curve.[2]

The second step executes the retention time alignment. This algorithm is feasible when considering two or more experiments with the same conditions (flowrate, column dimensions, etc.). In that case, it calculates the average time of peaks maxima and shifts all the data points to reach the lowest difference between average maxima values and actual maxima values of all

corresponding components.



*Form 1 – Measured data point and fitted asymmetric Gaussian model curves*

Optimization problem arises where is time of peak maximum for component *i* in experiment *j*. is average time of maxima of the peaks from all experiment for component *i*.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | *equation 3* |
|  |  | *equation 4* | |

The third step of the pre-processing workflow is the correction of feed duration according to total mass balance in the feed and in the end of the column. The total mass balance in the feed is expressed as

|  |  |  |
| --- | --- | --- |
|  |  | *equation 5* |

Where is concentration of component i in the feed and *d* is column diameter. The total mass at the end of the column is expressed as

|  |  |  |
| --- | --- | --- |
|  |  | *equation 6* |

Where is duration of the experiment. Subsequently, the optimization problem arises.

|  |  |  |
| --- | --- | --- |
|  |  | *equation 7* |

The determination of the *u*, *D* and is again optimization problem where in every step *equation 1* is numerically solved using implicit method of finite elements and the value of the sum of squared errors against measured data is calculated. The optimization problem was solved for each component separately. However, due to the shared parameter , the implementation of bilevel optimization is planned for following work.

Data pre-processing significantly reduced computing time of the final optimization algorithm and introduces higher consistency and reliability in EDM parameters determination process.

[1] DeVault D. (1943): The Theory of Chromatography. *Journal of the American Chemical Society 65*(4): 532-540.

[2] Green, M. and X. Chen (2020): Data Functionalization for Gas Chromatography in Python. *Journal of Chemical Education* 97(4): 1172-1175.

[3] Nicoud R.M. (2015): *Chromatographic Processes: Modeling, Simulation, and Design*.

Cambridge University Press.

[4] Rodrigues A.E., Pereira C., Minceva M. et al. (2015): *Simulated Moving Bed Technology*.

Butterworth-Heinemann, Oxford.