

Machine learning model for design of SMB processes and its application to separate rebaudioside A and stevioside

Chuanyi Yao^{1,2,*}, Zhenwei Zheng¹, Kuan Chen¹, Yinghua Lu^{1,2}, Shaokun Tang³, Enguo Fan^{4,*}

¹ Department of Chemical and Biochemical Engineering, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, 361005, China

² The Key Lab for Synthetic Biotechnology of Xiamen City, Xiamen University, Xiamen, 361005, China

³ Key Laboratory for Green Chemical Technology of Ministry of Education, School of Chemical Engineering & Technology, Tianjin University, Tianjin, 300072, China

⁴ Department of Microbiology and Parasitology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences/School of Basic Medicine, Peking Union Medical College, Beijing, China

Correspondence

Chuanyi Yao, Department of Chemical and Biochemical Engineering, College of Chemistry and Chemical Engineering, Xiamen University, 361005 Xiamen, Fujian, China

Email: cyao@xmu.edu.cn

Enguo Fan, Department of Microbiology and Parasitology, Institute of Basic Medical Sciences,

Chinese Academy of Medical Sciences/School of Basic Medicine, Peking Union Medical College,

Beijing, China

E-mail: enguo.fan@ibms.pumc.edu.cn

Abstract

Several machine learning algorithms were used to simulate the simulated moving bed (SMB) process, with the sugar separation of rebaudioside A and stevioside and enantioseparation of 1,1'-bi-2-naphthol racemate as case studies. It was found the random forest (RF) model and the deep neural network (DNN) model give satisfactory accuracy with MAEs lower than 0.19% (RF) and 0.08% (DNN). Then these two models were used to optimize the operation conditions for maximizing the feed flowrate under specific purity requirements. The RF model failed to give a set of operation conditions better than the training dataset. But the DNN model gave flowrates about 10% higher than the highest values in the training datasets, for both sugar separation and enantioseparation systems. Finally the optimized operation conditions for sugar separation were verified experimentally, with the final purities of rebaudioside A and stevioside being 99.2% and 98.8% respectively.

Keywords: Simulated moving bed, Random forests, Deep neural network, Rebaudioside A, Stevioside

1. Introduction

In the past few decades, machine learning (ML) has developed rapidly and provides an important assistance in numerous fields¹⁻³. Especially, the development of computer science has greatly improved the ability to train neural networks with large numbers of layers for hierarchical feature learning and thus, deep learning becomes a star ML technique because of its successful applications in speech recognition, image recognition, and natural language processing⁴. Nowadays, the application of ML in chemical engineering is also one of the rapid growing fields⁵, such as for process monitoring and online optimization^{6,7}, fault detection and diagnosis^{8,9}, construction of prediction models¹⁰⁻¹³, et al.

Simulated moving bed (SMB) technology is an important unit operation in chemical engineering to separate chemically similar compounds¹⁴⁻¹⁶. As a continuous multicolumn chromatographic process, SMB can achieve a higher productivity, higher purity, and higher yield while with a greatly reduced consumption of solvent and solid-phase comparing to the traditional preparative chromatography¹⁷⁻¹⁹. So it has been widely used in various fields, including fine petrol-chemicals, pharmaceuticals, and biotechnology industries²⁰⁻²². However, the design and optimization of SMB process is pretty complicated because of the coupling effects among operation conditions, including flow rates in different zones and switching periods, which makes a single factor optimization method not feasible. The usual requirement to simulate the dynamic processes of SMB separation is very time consuming^{23,24}, which needs to solve the mechanism model described by a series partial differential equations again and again until reaching a cyclic steady state, especially for optimization where the simulation processes at various operation conditions is repeated hundreds and even thousands of times.

Therefore, replacing the mechanism model with ML model has the potential to improve the efficiencies of design and optimization of SMB process^{23,25,26}.

Li et al.²⁷ applied the Bayesian regularization BP neural network model to fit a dataset that was generated from the mechanism model. The dataset contained 9730 groups of data with m_2 and m_3 as inputs and the concentration of the extract and raffinate as outputs. The simulation results showed that the deviation between the ML model and the mechanism model is small and the time required for the parameter optimization can be greatly decreased when using the ML model. Unfortunately, the optimization results are not verified experimentally.

In this work, various ML models are used and compared in order to replace the mechanism model of SMB process for the separation of rebaudioside A and stevioside. Rebaudioside A is a kind of natural sweeteners with a sweetness 150-300 times higher than that of sucrose²⁸. It is mainly produced by extraction from stevia (*Stevia rebaudiana*Bertonit). However, the crude extract usually contains certain amount of stevioside, which has a bitter taste²⁹. Because the high purity rebaudioside A has a much higher added value³⁰, the separation of rebaudioside A and stevioside received great attentions in recent years. As shown in Figure 1, the chemical structures of rebaudioside A and stevioside are very similar, which makes it difficult to achieve a high purity using macroporous adsorption resin^{31,32}. Although further utilization of the recrystallization method can obtain a high purity, the process is slow and complicated and thus limits the productivity and recovery³³. Therefore, it is the aim of our work to purify rebaudioside A through SMB for the first time. The ML models were compared first and two of them with satisfactory accuracy were then used for the optimization of operation conditions. The optimized conditions are verified by experiments. Furthermore, to prove the wide

applicability of ML models in SMB processes, the ML models were also used for another SMB system to separate chiral compounds with highly nonlinear isotherms because the isotherms of the studied system are linear.

2. Materials and methods

2.1 Materials and apparatus

Rebaudioside A (purity = 99.13%) and stevioside (purity = 99.55%) were purchased from Qufu Haigen stevia sugar Co., Ltd., Shandong. 1,3,5-tri-*tert*-butylbenzene (TTBB, purity = 98%) was purchased from Macklin Biochemical Co., Ltd, Shanghai. Acetonitrile (HPLC grade) were obtained from Cinc High Purily Solvents Co., Ltd., Shanghai. Pure water was provided by Shuibohui Co., Ltd., Xiamen.

The preparative columns with an inner diameter of 10 mm and a length of 150 mm were packed with 70 μm NH_2 adsorbent. The adsorbent was purchased from Beijing Greenherbs Science and Technology Development Co., Ltd. An analytical column (id. 4.6 mm \times 250 mm) packed with Supersil 5 μm NH_2 -S was purchased from Dalian Elite Analytical Instruments Co. Ltd.

A P230 II HPLC system (Dalian Elite Analytical Instruments Co. Ltd), comprising a P230 II pump and a UV230 II variable wavelength detector, was used to conduct the experiments, and EC2006 V1.80 software was used for data acquisition. The operation temperature was kept at 30.0°C and the detection wavelength was set at 210 nm.

A three-zone SMB system with a configuration of [1,1,2] was used to separate rebaudioside A and stevioside. Its schematic diagram is shown in Figure 2 and the device connection diagram for

SMB system can be found in our previous work³⁴.

2.2 Determination of porosity

The total porosity of the preparative column was determined by a non-retention indicator, 1,3,5-tri-tert-butylbenzene (TTBB). With NH₂ adsorbent as a stationary phase, pure acetonitrile as a mobile phase, column temperature setting at 30°C, the concentration of TTBB at 0.5 g/L and an injection volume of 20µL, the retention time (t_R) of TTBB in the preparative column was determined at various flowrates (Q). According to Eq. (1), t_R was used as the vertical coordinates and $1/Q$ as the abscissa, then the slope value, i.e. V_T , was obtained by linear fitting of experimental data.

$$t_R = \frac{V_T}{Q} \quad (1)$$

Where V_T is the sum of the total pore volume in the preparative column and the volume of the pipelines between injector and detector. After removing the preparative column, the retention time of TTBB was determined at a flowrate of 0.05 mL/min in order to determine the pipeline volume (V_L).

Then the total porosity was calculated by the Eq. (2):

$$\varepsilon_t = \frac{V_T - V_L}{V_c} \quad (2)$$

where V_c is the geometrical volume of the preparative column.

With the total porosity ε_t and inner porosity ε_p provided by the manufacturer, the external porosity ε_b can be obtained by the Eq.(3):

$$\varepsilon_t = \varepsilon_b + (1 - \varepsilon_b) \varepsilon_p \quad (3)$$

2.3 Determination of adsorption isotherms

The adsorption-desorption method is adopted to determine the adsorption isotherms of rebaudioside A and stevioside. First, the mobile phase containing a single component with known concentration (c_{mob}) was continuously delivered to the preparative column until the adsorption equilibrium was reached. Then, the column was washed with pure mobile phase to elute the component completely. The collected eluent was used to determine the component concentration in the eluent (c_{elu}). Through mass balance, the equilibrium concentration in the stationary phase can be calculated by Eq. (4):

$$q = \frac{c_{elu} V_{elu} - (\varepsilon_t V_c + V_L) c_{mob}}{(1 - \varepsilon_t) V_c} \quad (4)$$

where V_{elu} is the eluent volume.

2.4 Determination of diffusion and mass transfer coefficients

The axial diffusion coefficient (D_L) was estimated by the Chung and Wen correlation³⁵ :

$$D_L = \frac{\varepsilon_b d_p u}{0.2 + 0.011 (\Re)^{0.48}} \quad (5)$$

where d_p is the average diameter of adsorbent particles, u is the interstitial velocity of mobile phase,

and \Re is the Reynolds number.

The mass transfer coefficient of rebaudioside A ($k_{L,RA}$) and stevioside ($k_{L,ST}$) were both estimated by equation as follow³⁶:

$$\frac{1}{k_{e,i}} = H_i \left(\frac{d_p^2}{60 \varepsilon_p D_{p,i}} + \frac{d_p}{6 K_{f,i}} \right) \quad (6)$$

where H is the Henry constant, D_p is the intra-particle diffusion coefficient, and K_f is the film mass

transfer coefficient.

3. Results & discussion

3.1 Model parameters of SMB system

The transport-dispersive linear driving force model³⁷ was adopted to describe the dynamic behavior of rebaudioside A and stevioside in the columns. The model equations are presented as follows:

$$\frac{\partial c_i}{\partial t} + \frac{1-\varepsilon_b}{\varepsilon_b} \frac{\partial q_i}{\partial t} + u \frac{\partial c_i}{\partial x} = D_{L,i} \frac{\partial^2 c_i}{\partial x^2} \quad (7)$$

$$\frac{\partial q_i}{\partial t} = k_{e,i} (q_i^{\text{eq}} - q_i) \quad (8)$$

where c_i and q_i are the concentrations of component i in the liquid and stationary phase, $i = A$ (rebaudioside A) or B (stevioside), q_i^{eq} is the equilibrium concentration in the stationary phase, t is time

and x is the axial distance.

The initial conditions are:

$$t=0, c_i=q_i=0 \quad (9)$$

And the boundary conditions are:

$$x=0, \frac{\partial c_i}{\partial x} = \frac{u}{D_{L,i}} (c_i - c_i^{\text{eq}}) \quad (10)$$

$$x=L, \frac{\partial c_i}{\partial x} = 0 \quad (11)$$

where c_i^{eq} is the concentration of component i at the inlet of a column.

Eqs. (7-11) constitute a partial difference equation system, which was solved by the space-time conservation element/solution element method³⁸. Before solving the model equations, the model parameters including ε_b , D_L , k_e and adsorption isotherms of rebaudioside A and stevioside have to be determined first.

To determine the total porosity of the preparative column, the retention time of TTBB was measured at various flow rates, and the results are shown in Figure 3. Through linear regression of t_R vs $1/Q$, we obtained $V_T=8.748\text{mL}$. Additionally, the volume of the system pipelines was obtained as

$V_L = 0.007 \text{ mL}$. So the total porosity was calculated to be $\varepsilon_t = 0.742$. The internal porosity (ε_p) given

by the manufacturer was 0.563, then $\varepsilon_b = 0.41$ was obtained through Eq. (3).

The adsorption isotherms of rebaudioside A and stevioside were determined by the adsorption-desorption method, respectively. It was found that the equilibrium concentration (q^e) in the stationary phase has a linear relationship with the concentration (c) in the mobile phase (Figure 4) for both rebaudioside A and stevioside. By linear regression (with R^2 value of 0.9996 for rebaudioside A and 0.9995 for stevioside), the adsorption isotherms were obtained as follows:

$$q_A^e = 1.979 c_A \quad (12)$$

$$q_B^e = 1.581 c_B \quad (13)$$

Finally, the axial diffusion coefficient (D_L) and mass transfer coefficients were estimated by Eqs. (5-6). As shown in Figure 5, D_L varies with flow rate linearly. This is because the Reynolds number (\Re) in the denominator of Eq.(5) is so small in the range of experimental conditions that it has a little

effect on the axial diffusion coefficient. Through linear regression, the relationship between D_L and Q was obtained as Eq.(14).

$$D_L = 0.044 Q \quad (14)$$

The mass transfer coefficients of rebaudioside A and stevioside vary little with the flowrate (Figure 5). This is due to the slow internal diffusion rate and so the film transfer contributes little to the overall transfer coefficient³⁴. The average values with $k_{e,A}=21.27 \text{ min}^{-1}$ and $k_{e,B}=29.27 \text{ min}^{-1}$ are used in this work.

3.2 Dataset and preprocessing

To train a ML model for the prediction of the separation performance at certain operation conditions, a dataset with enough samples has to be constructed first. For the separation of rebaudioside A and stevioside by SMB system with a 3 zone SMB and a configuration of [1, 1, 2], the flow rate in Zone I (Q_I) was fixed at 2.0 mL/min to avoid that the pressure in the columns rises too high. Then the other three operation parameters including switching period (t_s) and flow rates in Zone II (Q_{II}) and Zone III (Q_{III}) were generated with the range and step size listed in Table 1. Obviously, the reasonable operation parameters in SMB separation must meet the conditions in Eq.(15):

$$Q_I, Q_{III} > Q_{II}, Q_{IV} \quad (15)$$

After filtering the samples violating Eq.(15), a dataset with 292,186 samples was obtained. The purities of extract and raffinate were used as the output after calculation by solving the model equation listed in Eqs.(7-11).

To avoid one term with a higher magnitude in the input parameters to dominate the ML model, the

Z-Score standardization³⁹ was used to normalize the input data, i.e. the input parameters were treated by Eq.(16):

$$x^{\hat{}} = \frac{x - \bar{x}}{\sigma} \quad (16)$$

where \bar{x} and σ is the mean and standard deviation of input parameter x . Finally, the dataset was randomly divided into two sets: a training set with 242,186 samples and a testing set with 50,000 samples. The training set was used to train different ML models and the selection of the hyperparameters, while the testing set was used to assess the performance of the final models.

Because the adsorption isotherms of rebaudioside A and stevioside are linear, to verify the wide applicability of ML model, the SMB system for enantioseparation of 1,1'-bi-2-naphthol racemate was selected to generate another dataset, which have highly nonlinear adsorption isotherms as shown in Eq.(17-18).

$$q_A^{\hat{}} = \frac{3.73 c_A}{1 + 0.0466 c_A + 0.0336 c_B} + \frac{0.3 c_A}{1 + 3 c_A + c_B} \quad (17)$$

$$q_B^{\hat{}} = \frac{2.69 c_B}{1 + 0.0466 c_A + 0.0336 c_B} + \frac{0.1 c_A}{1 + 3 c_A + c_B} \quad (18)$$

The parameters for this SMB process were obtained from literature^{40,41} and presented in Table S1 (supplemental material). The range and step size of input parameters to generate the dataset were listed in Table 1. After filtering the samples violating Eq.(15), a dataset with 262,753 samples was obtained. Similar to the sugar separation process, the input parameters of the dataset were normalized by Z-Score standardization. Then the dataset was randomly divided into two datasets with a testing dataset containing 50,000 samples and the rest samples as training dataset.

3.3 Comparison of different ML models

First, four typical ML models including support vector machines (SVM), k-nearest neighbors (KNN), decision trees (DT) and random forests (RF) were tested for the regression task of the two SMB separation processes. There are 3 input parameters for the sugar separation of rebaudioside A and stevioside and 4 input parameters for the enantioseparation of 1,1'-bi-2-naphthol racemate, respectively. While the output parameters are both 2 for the two SMB systems, i.e. the purities of extract (P_E) and raffinate (P_R). These algorithms were constructed by using Scikit-learn package⁴². The main hyperparameters, such as the kernel function in SVM, the k-value in KNN, the maximum depth in DT, and the number of estimators in RF were screened, and the other parameters remained as default values. With the mean absolute error (MAE) as the assessment criteria, the best results of these algorithms were listed in Table 2.

It can be seen from Table 2 that RF gave the best performance with the lowest MAEs of 0.114% for the sugar separation process and 0.185% for the enantioseparation process among the four ML models. The comparison of the target outputs and model predicted outputs by RF and the histograms of absolute error distribution for the sugar separation SMB systems were shown in Figure 6. It is obvious that the model predicted purities of extract and raffinate are in good agreement with the target values, with all the samples locating near the line of $y=x$. The MAE on the testing dataset is a little larger than that on the training dataset. This is reasonable because the model is obtained by learning the training dataset. A little difference between the MAEs on training and testing datasets indicates no overfitting occurs and the RF model has a good generalization ability.

The performance of RF model on the SMB system for enantioseparation of 1,1'-bi-2-naphthol racemate was shown in Figure 7. The model predicted the purities of extract and raffinate are also fit well with the target values. But the error is a little larger than that in sugar separation system. This may be due to the complexity incorporated by the highly nonlinear adsorption isotherms. In general, the MAE of lower than 0.2% has been smaller enough for application.

Although the RF model gave a satisfactory performance, we tried another model, deep neural network (DNN) to pursue a much higher accuracy. Because DNN is impressive in its high learning ability and has been successfully used in many fields for solving complex regression and classification tasks^{43–45}.

The structure of the network for the sugar separation SMB system was shown in Figure 8. The input layer contains 3 neurons corresponding to the operation conditions Q_{II} , Q_{III} and t_s . The output layer contains 2 neurons corresponding to P_E and P_R . The hidden layers contain several units, with each unit being composed of a dense layer and a batch normalization (BN) layer. BN layer was used to normalize the output of the last dense layer. This operation does not incorporate extra model parameters but has the ability to reduce the internal covariate shift and thus to accelerate training process⁴⁶. The rectified linear unit (ReLU) was selected as the activation function for the dense layer, while the sigmoid function was used as the activation function for the output layer to ensure that the predicted purities lie in the range of 0 to 1.

The TensorFlow platform⁴⁷ was used to build the DNN model. The mean square error (MSE) was selected as the loss function, and the adaptive moment estimation algorithm (Adam⁴⁸) was used as the optimizer with a learning rate of 0.0002. The dataset was fed to the training process with a batch size

of 128. The training was iterated to a maximum epochs of 6,000, which can be terminated earlier in case of the loss function on the testing dataset raised or unchanged for 500 epochs.

The number of neurons in each dense layer (referred to as width of the model hereafter) and the number of units in the hidden layers (referred to as depth of the model hereafter) are two important hyperparameters that would influence the model performance. So different combinations of the two values were tested and the results are listed in Table S2 (supplemental materials). While the width of the model is not too high, such as 64 and 128, the increase of depth will lead to the decrease of MAEs. And similarly, the increase of width at a fixed depth will also results in lower MAEs. This is reasonable because the increase of width and depth will make the model have a higher capacity and an improved learning ability. The model with a width of 256 and a depth of 4 gave an excellent performance with the MAEs of 0.0241% and 0.0708% for P_E and P_R , respectively. But a further increase of the width and depth will not bring significant improvement. So the DNN model with 4 units in the hidden layers and 256 neurons in each dense layer was selected as the appropriate model for the sugar separation process. The MAEs of P_E and P_R were 63% and 38% lower comparing to that of RF model, which is obviously a big improvement. Comparisons between the predicted purities and target values are shown in Fig 9. It can be seen that the DNN model fits excellent with the dataset and most of the points locate on the line of $y=x$.

For the enantioseparation process, the structure of the DNN model is similar to what is shown in Figure 8 except that the input layer contains 4 neurons corresponding to Q_{II} , Q_{III} , Q_{IV} and t_s . The MAEs at different combinations of width and depth of the model are listed in Table S3 (supplemental materials). It was found that the DNN model with a depth of 5 and a width of 512 gave a satisfactory

performance and so it was selected as the appropriate model. Comparing to the sugar separation process, the larger depth and width is understandable because of the increased complexity caused by the highly nonlinear isotherms in the enantioseparation process. The MAEs of P_E and P_R were 0.0295% and 0.0434%, respectively, and they are 84% and 77% lower compared to the values of RF model. The comparison between the predicted purities and target values shown in Figure 9 also underlined the high accuracy of DNN model.

It should be mentioned that, the improvement of the DNN model upon the RF model is much more significant for the enantioseparation process than that for the sugar separation process. For the RF model, the MAEs of enantioseparation process (0.180% for P_E and 0.185% for P_R) are much higher than that of sugar separation process (0.066% for P_E and 0.114% for P_R). This phenomenon is not found while using the DNN model, which may indicate that the DNN model has a stronger learning ability for the complex systems.

3.4 Optimization of SMB process using ML models

It is expected that the obtained ML models would be much more efficient in the simulation and optimization of SMB process. First the computational time for the simulation of the two SMB processes was tested using different models. All the evaluations were done on a personal computer with 3.00 GHz Inter core i5-8500 processor and 16 GB of RAM. The results are shown in Table 3. Compared to the mechanism model, the RF and DNN models can improve the simulation efficiency by four orders of magnitude.

The RF and DNN models were then used to optimize the sugar separation process in order to

maximize the feed flow rate (Q_F) while maintain the purities of extract and raffinate higher than 99.5%. The flow rate of zone I was fixed at 2.0 mL/min, and the decision variables were flowrates in zone II and zone III and switching period. The optimization problem is expressed as follows:

$$\begin{aligned} \max Q_F [Q_{II}, Q_{III}, t_s] \\ \text{subject to: } P_E, P_R \geq 99.5\% \end{aligned} \quad (19)$$

The BOX complex algorithm³⁴ was used to solve this optimization problem. While using the RF model to solve the purities at certain operation conditions, a maximum Q_F value of 0.120 mL/min was obtained. But for the DNN model, a higher Q_F value of 0.132 mL/min was obtained. It is interesting to compare the optimized results with the highest Q_F value at operation conditions that were included in the training dataset. After screening, it was found that the highest Q_F value meanwhile meeting the constraints of $P_E, P_R \geq 99.5\%$ is exactly 0.12 mL/min. This indicates that the RF model can not break through the information provided by the dataset, and so it is not appropriate for optimization. But DNN model has a higher generalization ability and it can give a set of operation conditions better than the whole training dataset. To verify this hypothesis, the SMB system for enantioseparation of 1,1'-bi-2-naphthol racemate was also optimized by use of the RF and DNN models. For this process, the flow rate in zone I was fixed at 56.83 mL/min and the constraints were $P_E \geq 93\%$ and $P_R \geq 96.2\%$. The optimized Q_F values were 6.5 and 7.19 mL/min for RF and DNN models,

respectively. While screening the training dataset, the highest Q_F that meets the purities constraints were found to be 6.5 mL/min. Again, the optimized Q_F value for RF model was not better than the

highest value in the dataset. But the DNN model resulted in a feed flowrate 10.6% higher than the best result in the training dataset. These results strengthen the hypothesis proposed above.

Finally, the optimized operation conditions using DNN model with $Q_{II}=1.66$ mL/min, $Q_{III}=1.792$ mL/min, and $t_s=10.09$ min, were tested on the laboratory apparatus for three times. The experimental concentrations of rebaudioside A and stevioside in the extract and raffinate with time are shown in Figure 10. After running 150 switching periods, the final purity of the extract and raffinate were assayed to be $99.2\pm0.3\%$ and $98.8\pm0.3\%$, respectively, which are in good agreement with the prediction of DNN model (both 99.5%).

4. Conclusion

Several machine learning models were tested to predict the product purity at certain operation conditions in the SMB process, with the sugar separation of rebaudioside A and stevioside as well as eantioseparation of 1,1'-bi-2-naphthol racemate as two case studies. The random forest and deep neural networks models gave satisfactory performances. The MAEs of RF and DNN models on the testing dataset were lower than 0.19% and 0.08%, respectively. However, while used for optimization, the RF model can not find a set of operation conditions better than the training dataset. DNN model gave feed flowrates about 10% higher than the highest value in the training dataset under specific purity requirements for the two case studies. The operation conditions for sugar separation obtained through optimization of DNN model were verified by experiments. The separation of rebaudioside A and stevioside was achieved by use of a four-column SMB system with purities of 99.2% and 98.8%

for rebaudioside A and stevioside, respectively. The simulation efficiency of DNN model is highly improved by about four orders of magnitude than the mechanism model. Such a high efficiency lays a foundation to apply machine learning models in industry for such as online simulation, optimization and control.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 21736009 & 22038012). Y. Yu and Z. Xu from Information and Network Center of Xiamen University are acknowledged for the help with the high-performance computing.

References

1. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans Pattern Anal Mach Intell.* 2018;40(4):834-848.
2. AlQuraishi M. ProteinNet: a standardized data set for machine learning of protein structure. *BMC Bioinformatics.* 2019;20:10.
3. Fu Y, Pan R, Yang Q, Gao W. Query-adaptive ranking with support vector machines for protein homology prediction. In: Chen J, Wang JX, Zelikovsky A, eds. *Bioinformatics Research and Applications.* Vol 6674. Springer-Verlag Berlin; 2011:320-331.
4. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436-444.
5. Lee JH, Shin J, Realff MJ. Machine learning: overview of the recent progresses and implications for the process systems engineering field. *Comput Chem Eng.* 2018;114:111-121.
6. Gonzaga JCB, Meleiro LAC, Kiang C, Maciel Filho R. ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process. *Comput Chem Eng.* 2009;33(1):43-49.
7. Li Z, Hong X, Hao K, Chen L, Huang B. Gaussian process regression with heteroscedastic noises —

- A machine-learning predictive variance approach. *Chem Eng Res Des.* 2020;157:162-173.
8. Md Nor N, Che Hassan CR, Hussain MA. A review of data-driven fault detection and diagnosis methods: applications in chemical process systems. *Rev Chem Eng.* 2020;36(4):513-553.
 9. Alauddin M, Khan F, Imtiaz S, Ahmed S. A bibliometric review and analysis of data-driven fault detection and diagnosis methods for process systems. *Ind Eng Chem Res.* 2018;57(32):10719-10735.
 10. Debiagi P, Nicolai H, Han W, Janicka J, Hasse C. Machine learning for predictive coal combustion CFD simulations-from detailed kinetics to HDMR reduced-order models. *Fuel.* 2020;274.
 11. Jirasek F, Alves RAS, Damay J, et al. Machine learning in thermodynamics: prediction of activity coefficients by matrix completion. *J Phys Chem Lett.* 2020;11(3):981-985.
 12. Hough BR, Beck DAC, Schwartz DT, Pfaendtner J. Application of machine learning to pyrolysis reaction networks: reducing model solution time to enable process optimization. *Comput Chem Eng.* 2017;104:56-63.
 13. Xu Y, Verma D, Sheridan RP, et al. Deep dive into machine learning models for protein engineering. *J Chem Inf Model.* 2020;60(6):2773-2790.
 14. Rajendran A, Paredes G, Mazzotti M. Simulated moving bed chromatography for the separation of enantiomers. *J Chromatogr A.* 2008;1216(4):709-738.
 15. Nogueira IBR, Martins MAF, Rodrigues AE, Loureiro JM, Ribeiro AM. Novel switch stabilizing model predictive control strategy applied in the control of a simulated moving bed for the separation of bi-naphthol enantiomers. *Ind Eng Chem Res.* 2020;59(5):1979-1988.
 16. Jiang X, Zhu L, Yu B, Su Q, Xu J, Yu W. Analyses of simulated moving bed with internal temperature gradients for binary separation of ketoprofen enantiomers using multi-objective optimization: Linear equilibria. *J Chromatogr A.* 2018;1531:131-142.
 17. He Q, von Lieres E, Sun Z, Zhao L. Model-based process design of a ternary protein separation using multi-step gradient ion-exchange SMB chromatography. *Comput Chem Eng.* 2020;138.
 18. Grill CA, Miller L, Yan TQ. Resolution of a racemic pharmaceutical intermediate: a comparison of preparative HPLC, steady state recycling, and simulated moving bed. *J Chromatogr A.* 2004;1026(1-2):101-108.
 19. Suvarov P, Lee JW, Wouwer A Vande, Seidel-Morgenstern A, Kienle A. Online estimation of

- optimal operating conditions for simulated moving bed chromatographic processes. *J Chromatogr A*. 2019;1602:266-272.
20. Wang J, Tian Y, Li Y, Xu J, Yu W, Ray AK. Multi-objective optimization of non-isothermal simulated moving bed reactor: methyl acetate synthesis. *Chem Eng J*. 2020;395.
 21. Mun S. Relationship between desorbent usage and the recovery of a target product in three-zone simulated moving bed processes designed under the conditions of positive and negative flow-rate-ratios of liquid to solid phases. *J Chromatogr A*. 2019;1603:388-395.
 22. Minceva M, Gomes PS, Meshko V, Rodrigues AE. Simulated moving bed reactor for isomerization and separation of p-xylene. *Chem Eng J*. 2008;140(1-3):305-323.
 23. Yan Z, Wang J, Wang S, Li S, Wang D, Sun W. Model predictive control method of simulated moving bed chromatographic separation process based on subspace system identification. *Math Probl Eng*. 2019;2019.
 24. Oh TH, Oh SK, Kim H, Lee K, Lee JM. Transition model for simulated moving bed under nonideal conditions. *Ind Eng Chem Res*. 2019;58(47):21625-21640.
 25. Hwang S, Song JR, Mai NL, Koo YM. Enhanced separation performance of three-zone simulated moving bed chromatography with ModiCon strategy. *KOREAN J Chem Eng*. 2020;37(6):1057-1065.
 26. He Q, Zhao L. Bayesian inference based process design and uncertainty analysis of simulated moving bed chromatographic systems. *Sep Purif Technol*. 2020;246:116856.
 27. Li L, Dai C, Deng J, Yuan D. Parameter optimization of neural network model based on simulated moving bed. In: *2019 Chinese Automation Congress (CAC)*. ; 2019:2506-2509.
 28. Hanson JR, Deoliveira BH. Stevioside and related sweet diterpenoid glycosides. *Nat Prod Rep*. 1993;10(3):301-309.
 29. Brandle JE, Starratt AN, Gijzen M. Stevia rebaudiana: its agricultural, biological, and chemical properties. *Can J PLANT Sci*. 1998;78(4):527-536.
 30. Xi D, Bhattacharjee J, Salazar-Gonzalez R-M, et al. Rebaudioside affords hepatoprotection ameliorating sugar sweetened beverage- induced nonalcoholic steatohepatitis. *Sci Rep*. 2020;10(1):6689.
 31. Liu Y, Di D, Bai Q, et al. Preparative separation and purification of rebaudioside A from steviol

- glycosides using mixed-mode macroporous adsorption resins. *J Agric Food Chem*. 2011;59(17):9629-9636.
32. Liu Y, Hua X, Wang M, Yang R. Purification of the mother liquor sugar from industrial stevia production through one-step adsorption by non-polar macroporous resin. *Food Chem*. 2019;274:337-344.
 33. Zhang Y, Chen T, Shi Z, He B. Studies on the separation of rebaudioside A by recrystallization. *Ion Exch Adsorpt*. 1998;14:515-520.
 34. Yao C, Chen J, Lu Y, Tang S, Fan E. Construction of an asynchronous three-zone simulated-moving-bed chromatography and its application for the separation of vanillin and syringaldehyde. *Chem Eng J*. 2018;331:644-651.
 35. Chung S, Wen C. Longitudinal dispersion of liquid flowing through fixed and fluidized beds. *AIChE J*. 1968;14(6):857-866.
 36. Kim TH, Yi SC, Suh YW, Lee KB, Mun S. Effect of ethanol content on mass-transfer parameters and hett indexes of amino ACIDS in a poly-4-vinylpyridine chromatography. *J Liq Chromatogr Relat Technol*. 2011;34(6):456-475.
 37. Yao H, Tian YC, Tade MO. Using wavelets for solving SMB separation process models. *Ind Eng Chem Res*. 2008;47(15):5585-5593.
 38. Yao C, Tang S, Lu Y, Yao H, Tade MO. Combination of space-time conservation element/solution element method and continuous prediction technique for accelerated simulation of simulated moving bed chromatography. *Chem Eng Process Intensif*. 2015;96:54-61.
 39. Chu CW, Holliday JD, Willett P. Effect of data standardization on chemical clustering and similarity searching. *J Chem Inf Model*. 2009;49(2):155-161.
 40. Pais LS, Loureiro JM, Rodrigues AE. Modeling, simulation and operation of a simulated moving bed for continuous chromatographic separation of 1,1'-bi-2-naphthol enantiomers. *J Chromatogr A*. 1997;769(1):25-35.
 41. Pais LS, Loureiro JM, Rodrigues AE. Modeling strategies for enantiomers separation by SMB chromatography. *AIChE J*. 1998;44(3):561-569.
 42. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn*

Res. 2012;12.

43. Ciaburro G, Iannace G, Passaro J, et al. Artificial neural network-based models for predicting the sound absorption coefficient of electrospun poly(vinyl pyrrolidone)/silica composite. *Appl Acoust.* 2020;169:107472.
44. Yao Z, Li J, Guan Z, Ye Y, Chen Y. Liver disease screening based on densely connected deep neural networks. *Neural Networks.* 2020;123:299-304.
45. Taghizadeh-Mehrjardi R, Mahdianpari M, Mohammadimanesh F, et al. Multi-task convolutional neural networks outperformed random forest for mapping soil particle size fractions in central Iran. *Geoderma.* 2020;376:114552.
46. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. Published online 2015.
47. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. Published online 2016.
48. Kingma D, Ba J. Adam: a method for stochastic optimization. *Int Conf Learn Represent.* Published online 2014.

Tables

Table 1. The range and step size of the input parameters to generate the dataset for sugar separation of rebaudioside A and stevioside and enantioseparation of 1,1'-bi-2-naphthol racemate.

Systems	Input parameter	Range	Step size	Unit
Sugar separation	Q_{II}	0.5 - 1.98	0.02	mL/min
	Q_{III}	0.52 - 2.50	0.02	mL/min
	t_s	5.0 - 15.05	0.15	min
Enantioseparation	Q_{II}	6.0 - 52.5	1.5	mL/min
	Q_{III}	8.0 - 54.5	1.5	mL/min
	Q_{IV}	6.0 - 52.5	1.5	mL/min
	t_s	2.0 - 4.3	0.1	min

Table 2. The MAEs of SVM, KNN, DT, and RF models for regression of the datasets generated by two SMB systems.

Systems	Outputs	SVM	NN	DT	RF
Sugar separation	P_E	5.175%	0.108%	0.304%	0.066%
	P_R	10.034%	0.439%	0.266%	0.114%
Enantioseparation	P_E	5.116%	0.507%	0.306%	0.180%
	P_R	7.012%	0.896%	0.265%	0.185%

Table 3. The computational time for simulation SMB process by using different models

SMB systems	RF model	DNN model	Mechanism model
Sugar separation	0.005 s	0.039 s	3 min
Enantioseparation	0.006	0.047	8 min

Figure Captions

Figure 1. The chemical structure of rebaudioside A and stevioside.

Figure 2. The schematic diagram of the four-column three-zone SMB with the configuration of [1,1,2]. The symbols D, E, F, and R stand for desorbent, extract, feed and raffinate, respectively.

Figure 3. The relationship between t_R of TTBB and $1/Q$ in preparative column at 30 °C.

Figure 4. The relationship between equilibrium concentration in stationary phase (q^e) and concentration in mobile phase (c) for rebaudioside A and stevioside.

Figure 5. The axial diffusion coefficient (D_L) and the mass transfer coefficients of rebaudioside A ($k_{e,RA}$) and stevioside ($k_{e,ST}$) at different flow rates.

Figure 6. Comparison of extract (a) and raffinate (b) purities predicted by random forest model and target values for the SMB separation of rebaudioside A and stevioside, and the distribution of absolute error of predicted extract purity (c) and raffinate purity (d) on the training and testing datasets.

Figure 7. Comparison of extract (a) and raffinate (b) purities predicted by random forest model and target values for the SMB separation of 1,1'-bi-2-naphthol racemate, and the distribution of absolute error of predicted extract purity (c) and raffinate purity (d) on the training and testing datasets.

Figure 8. The structure of the deep neural networks for modeling the SMB process for separation of rebaudioside A and stevioside.

Figure 9. Comparison of the purities of extract (a) and raffinate (b) in the separation of rebaudioside A and stevioside, and purities of extract (c) and raffinate (b) in the enantioseparation of 1,1'-bi-2-naphthol racemate with target values on the training and testing datasets.

Figure 10. The average concentrations of rebaudioside A (C_{EA} , C_{RA}) and stevioside (C_{EB} , C_{RB}) in the extract and raffinate, and the purities of extract (P_E) and raffinate (P_R) in a whole switching period at various switching numbers.