Uncertainty Quantification for Chromatography Model Parameters by Bayesian Inference using Sequential Monte Carlo Method

Yota Yamamoto, Tomoyuki Yajima, Yoshiaki Kawajiri*

*To whom correspondence should be addressed. Email: kawajiri@nagoya-u.jp
Dept. of Materials Process Engineering, Nagoya University, Nagoya, Japan

Abstract

Many model-based optimization methods have been proposed for chromatographic processes to ensure product quality and efficiency, but uncertainty of model parameters should be considered to assure robust design and operation. In this study, we developed a sequential Monte Carlo (SMC) parameter estimation method for chromatographic processes to estimate the parameter uncertainty rigorously within a reasonable amount of computation time. As an example, separation of glucose and fructose is considered. Through the example using artificial data, we confirmed that SMC can perform estimations more efficiently than the existing method, Markov chain Monte Carlo. Furthermore, through the example using lab-scale experimental data, we confirm that the time and effort for the sample analysis to identify the concentration of each component can be eliminated. We also examined the relationship between the number of cores and computation time for parallel implementation of SMC.

Keywords: Chromatography process, Parameter estimation, Uncertainty quantification, Bayesian inference, Sequential Monte Carlo

1. Introduction

Chromatography is one of the techniques for separation and purification. In its original form, it was developed for the separation of chlorophyll (Ettre and Sakodynskii, 1993), followed by improvements in adsorbents and solvents (Martin and Synge, 1941), and achieving higher flow rates (Horvath et al., 1967). In addition to the conventional batch chromatography, many advanced designs and operations have been developed and demonstrated, such as simulated moving bed (SMB) chromatography which allows continuous separation (Bentley and Kawajiri, 2013; Broughton and Gerhold, 1961; Francotte and Richert, 1997; Hashimoto et al., 1983; Kawajiri and Biegler, 2006).

Liquid chromatography is a system in which a feed solution containing multiple components flows through a column packed with adsorbents, and the components are separated by using the difference in the interaction between the components and the adsorbents. The adsorbent in the column is called the stationary phase, and the solvent of the solution flowing through the column is called the mobile phase (Guiochon et al., 2006). Figure 1 shows a schematic diagram of the separation of each component from a mixture of two components, A and B, using liquid chromatography. Figure 1 (a) - (d) show the time evolution of this process. Component A, represented by the circles, has stronger affinity towards the adsorbent, while component B, represented by the triangles, has weaker affinity towards the adsorbent. As shown in Figure 1 (a), the mixture of two components A and B dissolved in the mobile phase is pulsed into the column as feed, and then the mobile phase without the components continues to flow. The components are gradually separated by the difference in adsorption forces between the components and the adsorbent as they flow through the column. The mixture is separated and purified at the outlet of the column, as shown in Figure 1 (c) and (d).

Chromatography is widely used in industrial scales in chemical and pharmaceutical industries as a purification technique for the production of petrochemicals, sugars, and pharmaceuticals (Kaltenbrunner et al., 2007; Westerberg et al., 2012). Such large-scale processes require optimization of flow rates of solvents and feed, temperature, and choice of adsorbent to reduce manufacturing costs while meeting the required quality of the product. The more complex the process is, the more difficult it is to perform these operations relying on the experience of operators, which may lead to suboptimal design and operation, and failure in meeting requirements such as product purity. Therefore, it is desirable to employ model-based development, and estimation of parameters in the model is critical.

To obtain the parameters in the model, a common approach is to solve an inverse problem using experimental data and perform parameter estimation. If we only need a single set of estimated parameter values, least-squares methods can be used (Maeder and Zuberbuehler, 1990; Zhang, 1997). However, the result obtained by this method is a point estimate which does not quantify the uncertainty of the parameters. The uncertainty of the parameters, which may cause a model mismatch, must be

taken into account for robust design and operation of real processes, which is more challenging in complex processes such as SMB. To address this problem, there are many studies of methods for estimating parameter uncertainty in the past. As a frequentist method using the Fisher information matrix obtained from the Hessian of the objective function, a study of model comparison using lab-scale SMB (Grosfils et al., 2007), and insulin purification by single-column liquid chromatography (Borg et al., 2013) have been reported. However, these methods require assumptions such as approximating probability density functions to normal distributions. Such an approach does not necessarily guarantee the accuracy of the estimation (Froehlich et al., 2014; Joshi et al., 2006). Also, there are Bayesian methods using Markov Chain Monte Carlo (MCMC), such as a study of single-column drug separation (Briskot et al., 2019) and a study of lab-scale SMB protein separation (He and Zhao, 2020) to eliminate the shortcomings of the frequentist approaches. However, the method using MCMC requires a large number of serial sampling of parameters, which is computationally expensive (Conrad et al., 2018; Sherlock et al., 2017).

Uncertainty quantification (UQ) based on Bayesian inference, which utilizes prior knowledge, has been gaining attention in recent years. To deal with complex models for chemical engineering applications, Markov Chain Monte Carlo (MCMC) has been explored (Na et al. 2019; Hsu et al. 2009). This algorithm has been used also for single-column drug separation (Briskot et al., 2019) and lab-scale SMB protein separation (He and Zhao, 2020). However, MCMC requires a large number of serial sampling of parameters, which is computationally expensive (Conrad et al., 2018; Sherlock et al., 2017). To reduce the computational effort, alternative algorithms have been proposed, such as Hamiltonian Monte Carlo and Sequential Monte Carlo (SMC) (Barbu and Zhu, 2020). Applications of SMC to state estimation with *data tempering*, or particle filtering, can be found in literature (Rawlings and Bakshi, 2006). Recently, the SMC algorithm has been applied also to parameter estimation of gas adsorption processes with *likelihood tempering* (Kalnayanaraman et al., 2016). Nevertheless, this approach has yet to be applied to chromatographic processes.

In this study, we present sequential Monte Carlo (SMC) parameter estimation and uncertainty quantification (UQ) for single-column liquid chromatography. The SMC applied in this study employs *likelihood tempering* (Kalyanaraman et al., 2016), in which the prior distribution reflects the likelihood in a stepwise manner as it approaches the posterior distribution. By implementing parallel computing, Bayesian inference on parameters can be performed in a shorter amount of time. As an example, we adopted the process of separating glucose and fructose by lab-scale cation exchange chromatography, considering artificial data obtained by simulation, and actual lab-scale experimental data. Furthermore, the SMC estimation was compared to that by MCMC. We also demonstrate that experimental effort for component analysis may be eliminated by validating the *numerical deconvolution* of two components from total concentrations utilizing the model with the estimated parameters. Finally, the relationship between the number of cores used for parallel computation and the computation time for

parameter estimation is also examined.

The remainder of this paper is as follows: Section 2 describes the background, including the mathematical model of the target process and Bayesian inference techniques; Section 3 describes the specific experimental and computational methods; Section 4 presents the estimation results and discussions; and finally, Section 5 is the conclusion.

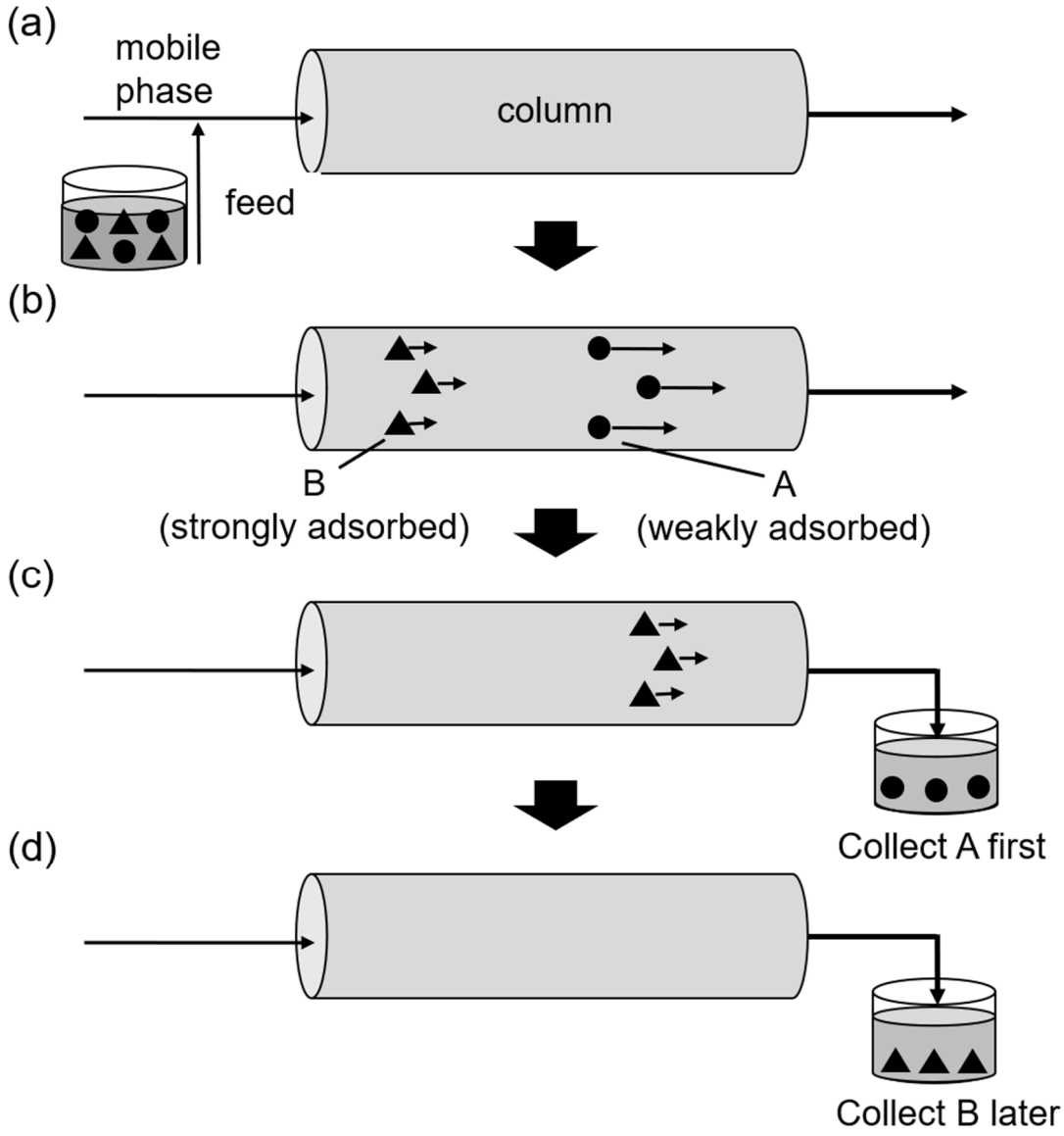

Figure 1. Schematic diagram of two-component separation by liquid chromatography

2. Background and methods

2.1 Experimental and estimation methods for model parameters

To model chromatographic processes, some tests are carried out using a column, such as pulse injection and step injection (breakthrough) tests. In a pulse injection test, a small amount of the feed mixture is injected at the inlet of the column, and the concentrations at the outlet are measured to record chromatograms (Figure 1). Since the injected volume is usually small, the concentrations in the column are low. In this test, information on the adsorption equilibria and mass transfer kinetics at low concentrations can be obtained. On the other hand, in a step injection test, the feed mixture is supplied continuously at the column inlet to record breakthrough curves. In this test, the concentrations in the column remain high, and thus information on the adsorption equilibria at high concentrations can be obtained. It is known that step injection tests are effective for a system described by nonlinear isotherms, where the equilibrium behavior is dependent on the concentrations (Guiochon et al., 2006).

From the chromatograms and breakthrough curves obtained from the above tests, parameters in the model equations can be determined. A common approach is to use correlations of the model parameters and moments of the chromatograms and breakthrough curves (Guiochon et al., 2006). An alternative approach is regression, or the inverse method, where a mathematical model is fitted to the chromatograms and breakthrough curves to identify a single set of the optimal parameters (Grosfils et al., 2007).

In this work, we explore an alternative approach of sampling based on Bayesian inference from the chromatograms and breakthrough curves. Unlike the regression approach where only a single set of model parameter values is pursued, the sampling approach obtain the probability distributions of the model parameters. The probability distributions quantify the uncertainty of the model parameters, which will be critical information for robust design and operation. In this study, we implement two sampling techniques, MCMC and SMC, as described in Section 2.4 and 2.5.

2.2. Numerical model

Among many numerical models proposed for liquid chromatography (Kawajiri, 2021; Schmidt-Traub et al., 2012), in this study, we used the linear driving force (LDF) model (Grosfils et al., 2007; Kawajiri and Biegler, 2006; Moon and Kook Lee, 1986; Wang and Tien, 1982) which consists of two mass balance equations and one adsorption isotherm equation. The LDF model assumes that the effects of heat and pressure in the column and diffusion in the radial direction can be neglected, and that the diffusion in the axial direction and the pores of the adsorbent can be described by mass transfer between the liquid and solid phases. These assumptions are commonly used in literature (Grosfils et al., 2007; Guiochon et al., 2006).

The equations that make up the LDF model are as follows. First, the mass balance equation in the liquid phase is given by:

$$\frac{\partial C_i(x,t)}{\partial t} = -\frac{1-\varepsilon_b}{\varepsilon_b}\frac{\partial q_i(x,t)}{\partial t} - \frac{u}{\varepsilon_b}\frac{\partial C_i(x,t)}{\partial x}, \tag{1}$$

where the subscript i denotes the adsorbent components, which in this study are glucose (Glu) and fructose (Fru); C_i is the concentration of component i in the liquid phase; q_i is the concentration of component i in the solid phase; ε_b is the overall porosity in the column; u is the superficial velocity of the liquid; t is the time; and x is the coordinate in the axial direction, where the inlet of the column is $x = 0$.

The mass balance equation in the solid phase is given by:

$$\frac{\partial q_i(x,t)}{\partial t} = \frac{K_i}{1-\varepsilon_b}\left(C_i(x,t) - C_i^{eq}(x,t)\right), \tag{2}$$

where K_i is the overall mass transfer coefficient of component i based on the liquid phase; and C_i^{eq} is the equilibrium concentration of component i in the liquid phase given by the adsorption equilibrium.

The initial and boundary conditions for the partial differential equations (PDEs), (1) and (2), are given as follows. The initial conditions are given by:

$$C_i(x,0) = q_i(x,0) = 0, \tag{3}$$

and the boundary conditions are given by:

$$C_i(0,t) = C_{i,in}, \tag{4}$$

where subscript in represents the inflow to the column, and $C_{i,in}$ is the concentration of component i in the liquid phase flowing into the column. In this study, the feed is injected by pulse or step input, and $C_{i,in}$ is given by:

$$C_{i,in} = \begin{cases} C_{i,feed}, & 0 \le t \le t_f \\ 0, & t < 0, t_f < t, \end{cases} \tag{5}$$

where $C_{i,feed}$ is the concentration of component i in the feed; and t_f is the injection time of the feed where for the step input $t_f = \infty$.

The adsorption isotherm equation for the adsorption equilibrium between the liquid and solid phases is given by:

$$q_i^{eq}(x,t) = \frac{H_i C_i(x,t)}{1 - b_{Glu}C_{Fru}(x,t) - b_{Glu}C_{Glu}(x,t)}, \tag{6}$$

where q_i^{eq} is the concentration of component i in the solid phase at the adsorption equilibrium, H_i is Henry's constant for component i, and b_i is the adsorption equilibrium constant for component i. Equation (6) is called anti-Langmuir isotherm. Although many models have been proposed to describe the adsorption isotherm (Guiochon et al., 2006), previous studies (Chilamkurthi et al., 2012; Nowak et al., 2007) have shown that the adsorption equilibrium between sugars and cation exchange resin, which is demonstrated in this study, can be well described by the anti-Langmuir isotherm equation. This model assumes that the equilibrium concentration in the solid phase increases as the

concentration in the liquid phase increases.

2.3. Bayesian inference

Bayesian inference is the process of inferring the causal phenomena (parameters) from observed facts (data) as probability distributions based on Bayes' theorem (Gelman et al., 2013). Bayes' theorem is given by:

$$P(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{P(\boldsymbol{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\boldsymbol{y})}, \tag{7}$$

where in general, $\boldsymbol{y}$ are the observed data, and $\boldsymbol{\theta}$ are the parameters. In this study, $\boldsymbol{y} = \{(t_j, \boldsymbol{C}_j)\}_{j=1}^{M}$ are observed data obtained by experiments as a set of discrete-time points and concentrations of components at the column outlet, and $\boldsymbol{\theta} = [H_{Glu}, H_{Fru}, K_{Glu}, K_{Fru}, b_{Glu}, b_{Fru}, \sigma]^{\mathrm{T}}$ are model parameters to be estimated, M is the number of observations of the data, and σ is the standard deviation assuming that the observation errors of all data follow a constant normal distribution. In the above equation, $P(\boldsymbol{\theta}|\boldsymbol{y})$ is the posterior distribution, $P(\boldsymbol{y}|\boldsymbol{\theta})$ is the likelihood distribution, $P(\boldsymbol{\theta})$ is the prior distribution, and $P(\boldsymbol{y})$ is the marginal likelihood.

The likelihood distribution is given by:

$$P(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{j=1}^{M} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{\boldsymbol{C}_j{}^{model} - \boldsymbol{C}_j{}^{data}}{\sigma}\right)^2\right\}, \tag{8}$$

when the observation error follows a normal distribution with a standard deviation σ. Here, the superscript $model$ refers to the simulation data computed by the LDF model, and the subscript $data$ refers to the experimental data.

The prior distribution is the parameter distribution before taking the observed data into account. If there is no prior knowledge of the parameters at all, a uniform distribution without any restrictions can be set as an uninformative prior distribution. On the other hand, if the estimator has some prior knowledge of the parameters such as relevant data, empirical rules, etc., it is also possible to set a uniform distribution with upper and lower limits or a normal distribution with the estimated value by maximum likelihood as the mean as the prior distribution. The parameters by maximum likelihood estimation are given by:

$$\boldsymbol{\theta}^{mle} = arg\max_{\boldsymbol{\theta}} P(\boldsymbol{y}|\boldsymbol{\theta}) = arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{M} \left(\frac{\boldsymbol{C}_j{}^{model} - \boldsymbol{C}_j{}^{data}}{\sigma}\right)^2 + 2M \cdot \log(\sigma), \tag{9}$$

where the subscript mle indicates an estimate from maximum likelihood estimation. The latter equality in Equation (9) is given because the negative log-likelihood is minimized, which shows that

the result of maximum likelihood estimation is equal to the least square minimization when the observation error follows a normal distribution with a constant value of σ.

The marginal likelihood $P(\boldsymbol{y})$ refers to the probability distribution of obtaining the observed data *y*. From the law of total probability, it is given by:

$$P(\boldsymbol{y}) = \int P(\boldsymbol{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})d\boldsymbol{\theta}\,. \tag{10}$$

However, this integral is difficult to solve either analytically or numerically. Therefore, it is not possible to calculate the equation (10) directly to obtain the posterior distribution. One approach to obtain the posterior distribution is random sampling following the posterior distribution, using the information of the prior and likelihood distributions.

2.4. Markov Chain Monte Carlo (MCMC)

MCMC (Gelman et al., 2013) is a general term for random number sampling methods using Markov chains, which can be used to sample values according to the probability density of the posterior distribution. The posterior distribution can be approximated by a sufficiently large number of samples without fixing a functional form of probability distributions. Various algorithms for MCMC have been developed, which may result in different computational load and convergence properties. The Metropolis-Hastings (M-H) algorithm (Hastings, 1970) is one of the popular methods and is often used as a basic and general-purpose MCMC. The outline of MCMC using M-H algorithm is as follows.

1. Initialization and determination of the proposal distribution

   Choose the initial values of the parameters $\boldsymbol{\theta}_0$ to start sampling and the proposal distribution $\boldsymbol{Q}(\boldsymbol{\theta}^*|\boldsymbol{\theta}_k)$ to propose parameter updates, where $\boldsymbol{\theta}^*$ is the candidate parameter and $\boldsymbol{\theta}_k$ is the parameter adopted in the kth sampling. Here, $\boldsymbol{\theta}_0$ can be set to any arbitrary value within the range of the prior distribution, and the maximum a posteriori (MAP) estimate, $\boldsymbol{\theta}^{MAP}$, may be employed to speed up the convergence of MCMC. Note that when a sufficiently wide uniform distribution is used in the MAP estimation, then $\boldsymbol{\theta}^{MAP} \cong \boldsymbol{\theta}^{mle}$.

2. Update parameters for $k \to k+1$

   Based on the current values of the parameters $\boldsymbol{\theta}_k$, candidate parameters $\boldsymbol{\theta}^*$are generated from the proposal distribution as follows:

   $$\boldsymbol{\theta}^* \sim \boldsymbol{Q}(\boldsymbol{\theta}^*|\boldsymbol{\theta}_k). \tag{11}$$

   Using $\boldsymbol{\theta}^*$, we calculate the adoption rate, α, which can be calculated using Bayes' theorem as follows:

   $$\alpha = \frac{P(\boldsymbol{\theta}^*|\boldsymbol{y})\boldsymbol{Q}(\boldsymbol{\theta}_k|\boldsymbol{\theta}^*)}{P(\boldsymbol{\theta}_k|\boldsymbol{y})\boldsymbol{Q}(\boldsymbol{\theta}^*|\boldsymbol{\theta}_k)} = \frac{P(\boldsymbol{y}|\boldsymbol{\theta}^*)P(\boldsymbol{\theta}^*)\boldsymbol{Q}(\boldsymbol{\theta}_k|\boldsymbol{\theta}^*)}{P(\boldsymbol{y}|\boldsymbol{\theta}_k)P(\boldsymbol{\theta}_k)\boldsymbol{Q}(\boldsymbol{\theta}^*|\boldsymbol{\theta}_k)}. \tag{12}$$

   Using α, the adoption or rejection of $\boldsymbol{\theta}^*$ is determined probabilistically as follows:

$$\boldsymbol{\theta}_{k+1} = \begin{cases} \boldsymbol{\theta}^*, & \beta \leq \alpha \\ \boldsymbol{\theta}_k, & \beta > \alpha, \end{cases} \tag{13}$$

where β is a uniform random number generated as $\beta \sim U(0,1)$. This series of parameter updates are repeated until the parameter distributions converge.

A special case of M-H algorithm where the proposal distribution is symmetric is also called the Metropolis algorithm (Metropolis et al., 1953). By sampling a sufficient number of times with the M-H algorithm, a set of numbers proportional to the target posterior distribution can be obtained. To eliminate the influence of the choice of the initial value $\boldsymbol{\theta}_0$ in the early stages of sampling, the initial sampling period is truncated, which is called burn-in. MCMC has been used in many studies to perform complex Bayesian inference, but it tends to require high computational cost to carry out a large number of samplings. In particular, since the numerical model used in this study includes PDEs, the computational cost of a single sampling (solution of the PDE model) is very high. Therefore, it is critical to employ an efficient algorithm.

2.5. Sequential Monte Carlo (SMC)

SMC (Gordon et al., 1993; Kitagawa, 1996) is a method for estimating the posterior distribution by having a large number of particles track variables (in this study, they are parameters) using information from observed data. It was originally developed for sequential estimation of state variables in state-space models and is now widely used for forecasting time series data (Lang et al., 2007). SMC consists of four major phases: initialization, mutation, likelihood calculation, and resampling. It can be used to estimate quantities including state variables and model parameters by improving mutation and likelihood computation (Gao and Zhang, 2012; Kalyanaraman et al., 2016; Kantas et al., 2009). For example, estimation using self-organizing state-space models, in which parameters are considered to be part of the state variables, and data tempering, in which the number of observed data is gradually added to anneal the prior distribution to the posterior distribution, have been studied.

In this study, we adopted *likelihood tempering* (Kalyanaraman et al., 2016; Uppsala University, 2017), which brings the parameter distributions closer to the posterior distributions by gradually strengthening the influence of the likelihood distribution during the likelihood calculation phase. The distribution of the parameters in the likelihood tempering is given as follows:

$$\pi_l(\boldsymbol{\theta}) = P(\boldsymbol{y}|\boldsymbol{\theta})^{\gamma_l} P(\boldsymbol{\theta}). \tag{14}$$

Here, $l = 0,1,\cdots,L$ are the steps repeated for a total of L times, π_l is the distribution of the parameters at the lth step, and γ_l is a monotonically increasing scalar value satisfying $\gamma_0 = 0$ and $\gamma_L = 1$; that is, $\pi_0(\boldsymbol{\theta}) = P(\boldsymbol{\theta})$ at $l = 0$ and $\pi_L(\boldsymbol{\theta}) = P(\boldsymbol{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})$ at $l = L$. The sequence shown below gives the Bayesian inference:

1. Initialization at $l = 0$

Determine the number of particles, N_P, to be computed. For $m = 1,2,\cdots,N_P$, we generate particles, $\boldsymbol{\theta}_0^m \sim P(\boldsymbol{\theta})$, that follow the prior distribution. Let the weight of the particle $\boldsymbol{\theta}_0^m$ be $w_0^m = 1/N_P$; that is, the weight of each particle is equal and the sum is 1.

2. Mutation for $l-1 \to l$

   Mutate each particle using the MCMC step; that is, consider $k = l-1$ in equations (11), (12), and (13), and calculate $\boldsymbol{\theta}_l^m$ from $\boldsymbol{\theta}_{l-1}^m$. In this work, we chose the normal distribution with the covariance matrix $\boldsymbol{\Sigma}_k$ for proposal distribution ($\boldsymbol{Q} = \boldsymbol{N}(\boldsymbol{\theta}_k, \boldsymbol{\Sigma}_k)$).

3. Likelihood calculation

   Calculate the likelihood of the particles at each step of the likelihood tempering as normalized weights as follows:

$$\widetilde{w}_l^m = \frac{w_{l-1}^m P(\boldsymbol{y}|\boldsymbol{\theta}_l^m)^{\gamma_l - \gamma_{l-1}}}{\sum_{m=1}^{N_P}\{w_{l-1}^m P(\boldsymbol{y}|\boldsymbol{\theta}_l^m)^{\gamma_l - \gamma_{l-1}}\}}. \tag{15}$$

4. Resampling

   Redistribute particles proportionally to the value of $\widetilde{w}_l^m$; that is, if $\widetilde{w}_l^m$ is large, particles are replicated; conversely, if $\widetilde{w}_l^m$ is too small, the particles are annihilated. In this work, the number of particles for each parameter value, which is an integer, is chosen to be nearly $N_P \widetilde{w}_l^m$. After this particle redistribution, the weight for each particle is renormalized equally so that $w_l^m = 1/N_P$.

Figure 2 illustrates the sequence of the SMC algorithm (Leeuwen, 2009). Since SMC is computed by spreading many particles widely, it converges to the posterior distribution faster than MCMC, which starts from a single particle. We also note that since the particles do not interfere with each other (each particle is independent), the calculation can be performed in parallel, which allows utilizing multiple cores simultaneously.

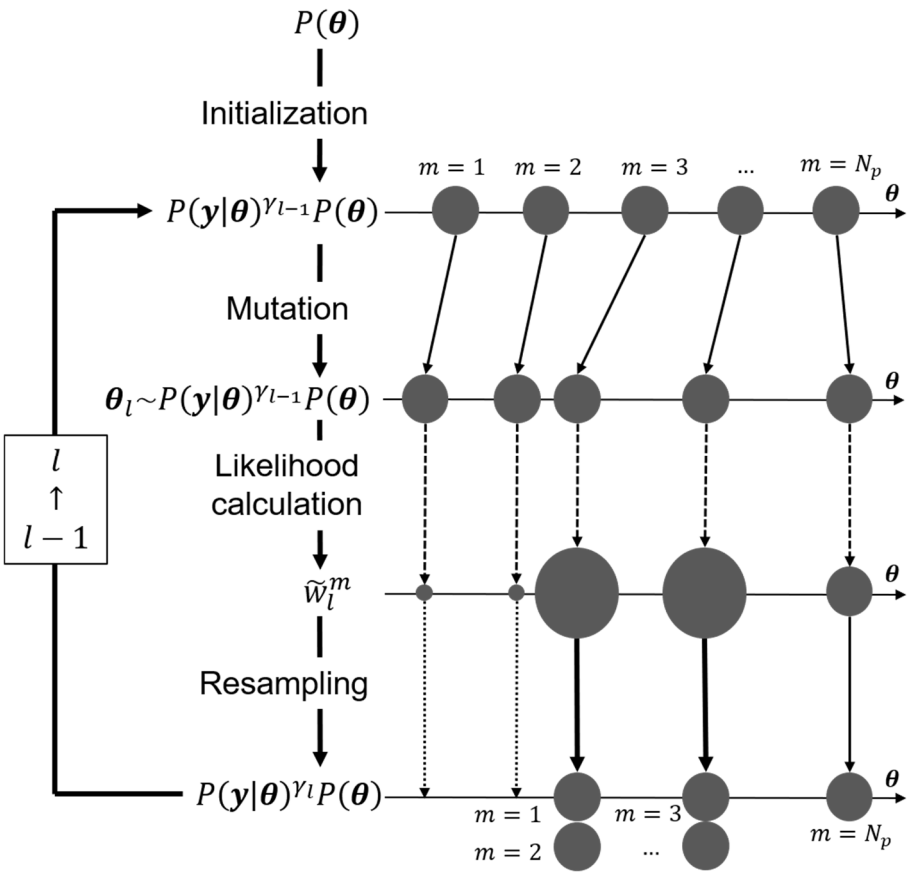

Figure 2. Schematic diagram of SMC

3. Experimental and Computational conditions

3.1. Case.1: Simulated experiments to generate artificial data

In the simulated experiments, the numerical model (1)-(6) with all parameters specified in advance was solved in Python. The outlet concentrations from the column in this simulation were added to artificial random error that follows a normal distribution to obtain the artificial data, $\boldsymbol{y}$. The parameters related to the HPLC system were set according to the values used in the lab-scale experiments, as shown in Table 1. In this table, three different feed injection conditions, (A), (B), and (C) in the experiments and the true values of parameters to be estimated are also shown. These parameters were chosen referring to previous experimental reports on the separation of sugars (Nowak et al., 2007; Sreedhar and Kawajiri, 2014).

3.2. Case.2: Lab-scale experiments

In the lab-scale experiments, pulse and step injection tests of glucose and fructose by column chromatography were conducted using an actual HPLC system. The HPLC system consisting of degasser DGU-20A3R, pump LC-20AD, autosampler SIL-20AHT, column oven CTO-20A, UV/VIS detector SPD-20A, RI detector RID-20A, and system controller CBM-20A was used for all experiments which was operated through the control software LCsolution (all of the above SHIMADZU CORPORATION, Kyoto, Japan). A schematic diagram of the HPLC system is shown in supplementary material (Figure S1). First, the dead volume was measured using dextran 40,000 (FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) as an indicator with the column removed, estimated to be 0.259 mL. Next, the overall porosity of the preparative column was measured also using the dextran, which was estimated to be 0.397. For the preparative column, an industrial cation exchange resin DIAION^{TM} UBK535 (Mitsubishi Chemical Corporation, Tokyo, Japan) was packed in a custom-made empty stainless column (Sugiyama Shoji Co., Ltd., Kanagawa, Japan).

Using the column, two pulse injection tests and one step injection test were conducted under the same conditions as in (A), (B), and (C) of the simulated experiment, using aqueous solutions of D(+)-glucose and D(-)-fructose (both FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) as feed components. The pulse input in (A) and (B) were performed by the autosampler, and the step input was performed using a manual injector with a sufficiently large sample loop (20mL). Pure water was used as the mobile phase.

The data sets for parameter estimation were obtained as follows. The feed injections were done under each of the experimental conditions shown in Table 1. At the outlet from the column, the eluent is supplied to the RI detector. The data were obtained by converting the intensity values of the RI detector into concentrations using values of feed concentration, inject volume, peak area, and flow rate.

In addition to the data sets obtained by the RI detector in (A), (B), and (C), an additional data set was obtained using another measurement technique only in (C), which is used for the validation test in Section 4.1.1 and 4.2.1. For the large-volume injection test, (C), the solution at the outlet was fractionated at an interval of 25 s, and the concentration of each sample was analyzed using an analytical HPLC column, SUGAR SC1011 (Showa Denko K.K., Tokyo, Japan). In this analysis, the concentration was converted from the RI intensity using the calibration curves prepared from standard solutions of glucose and fructose prepared separately (0.500 g/L, 1.00 g/L, 2.00 g/L, 5.00 g/L, 10.0 g/L, 20.0 g/L, 30.0 g/L, and 50.0 g/L). The flow rate was of the analysis was 1.00 mL/min and the temperature of the column oven was 80 ℃. It should be noted that obtaining this data set requires experimental effort, and we attempt to eliminate the need to carry out such measurement. The experimental conditions are summarized in Table 1, and the data sets in our study are summarized in Table 2.

Table 1. Conditions in obtaining experimental data

Experiment	Simulated			Lab-scale			
Inject conditions	(A)	(B)	(C)	(A)	(B)	(C)	Units
Column length	250						mm
Column diameter	10.0						mm
Flow rate	1.00						mL/min
Temperature	–			60			℃
Column porosity	0.397						–
Feed concentration of Glucose	250	0	250	250	0	250	g/L
Feed concentration of Fructose	0	250	250	0	250	250	g/L
Input type	Pulse	Pulse	Step	Pulse	Pulse	Step	-
Injection volume	0.100	0.100	–	0.100	0.100	–	mL
Adsorption resin	-			DIAION™ UBK535			-
True value of H_{Glu}	0.301			Unknown			–
True value of H_{Fru}	0.531			Unknown			–
True value of K_{Glu}	4.70×10^{-3}			Unknown			s^{-1}
True value of K_{Fru}	8.30×10^{-3}			Unknown			s^{-1}
True value of b_{Glu}	6.34×10^{-4}			Unknown			L/g
True value of b_{Fru}	2.48×10^{-4}			Unknown			L/g
True value of σ	5.00×10^{-3}			Unknown			–

Table 2. Summary of data sets.

	Data for estimation	Data for validation test (experimental deconvolution data)
Experiment	(A), (B), (C)	(C)
Data frequency	Online and continuous*	Sampling discretely at 25 sec intervals
Experimental measurement technique	Online detector *	HPLC analysis of each sample
Data description	Overall (total) concentration	Individual concentrations of the two components

*In the lab experiments in our study, data in (C) were obtained using HPLC analysis and fractionating the column effluent instead of online RI detection to avoid the saturation of the RI signal.

3.3. Computational environment and numerical methods

3.3.1. Solution of the numerical model

The solution of the system of PDEs consisting of equations (1)-(6) was performed by the numerical method of lines (Schiesser, 1991), where the PDE system was discretized in the spatial domain. The resulting set of ordinary differential equations (ODEs) which includes only temporal derivatives was solved numerically. The spatial discretization was performed using the central difference method with 100 finite differences. The ODE system was solved numerically using *scipy.integrate.odeint* and the integration algorithm was *lsoda*. To speed up the computation, we used *numba*, a JIT compiler for Python (Lam et al., 2015).

All computations, including simulations, MCMC, and SMC, were implemented in Python. The computing environment was an Intel(R) Core(TM)i9-10940X CPU at 3.30 GHz with 14 cores (28 logical processors).

3.3.2. Execution conditions for MCMC and SMC

First, MAP solutions for each parameter were obtained, where uniform distributions were employed as the prior distributions. The range of Henry's constants H_i and overall mass transfer coefficients K_i were set to ± 50 % of the initial value estimated by the initial estimation method (Guiochon et al., 2006), and adsorption equilibrium constants b_i and standard deviations σ were set to be sufficiently wide in the positive domain to cover all potential values.

Using the MAP solutions, the prior distributions for UQ of parameters by MCMC and SMC were determined. For Henry's constants H_i, overall mass transfer coefficients K_i, and the standard deviation σ, the prior distributions are normal distributions where the mean is $\boldsymbol{\theta}^{MAP}$ with sufficiently wide standard deviations; standard deviation for σ is $0.1 \times \sigma^{MAP}$, those for H_i and K_i are $0.05 \times H_i{}^{MAP}$ and $0.05 \times K_i{}^{MAP}$, respectively. For adsorption equilibrium constants b_i, we set a uniform distribution with the same width as the prior distribution for MAP estimation because of its low sensitivity and difficulty in estimation.

The data used in the estimation was prepared as follows. All data of outlet concentrations used in the likelihood calculations were treated as dimensionless values by dividing them by the maximum concentrations. This was done to avoid overfitting the data with high concentrations when dealing with multiple data sets. Three types of data were used for each of simulated experiments and lab-scale experiments: time-series data of glucose concentration obtained in experiment (A), time-series data of fructose concentration obtained in experiment (B), and time-series data of total concentration of

glucose and fructose obtained in experiment (C).

The MCMC calculations were performed using *PyMC*, a MCMC library for Python. The total number of samples was 50,000, and the burn-in was 30,000, and samples were thinned out every other one to eliminate the effect of autocorrelation between samples. Convergence was judged by visual inspection of the trace plots.

The SMC calculations were performed using a program developed by the authors (Yamamoto, 2021). The number of particles was 10,000, and the number of steps for likelihood tempering L was 24. After the likelihood tempering was completed, only the mutation step was iterated five times additionally to confirm the convergence of the posterior distributions to confirm that the shapes of the distributions do not change. The SMC was run on a workstation that has 14 cores using *joblib*, a parallel computing library for Python. The prior distributions for the artificial and lab-scale experimental data are shown in supplementary material, and the hyperparameters used in the calculations are also shown in that.

4. Results and Discussions

4.1. Case.1: Simulated experiment

The purpose of this case study is to confirm that the proposed approach finds the correct values of the parameters, and the SMC and MCMC find the same estimates. A density plot matrix of the posterior distributions of the parameters obtained by SMC is shown in Figure 3. Each of the histograms on the diagonal represents the posterior probability distribution of each parameter itself. The off-diagonal density plots represent the correlation between the two parameters. We can see that there is a strong negative correlation between b_{Glu} and b_{Fru}. This can be attributed to the fact that these two parameters have nearly the same influence on the denominator value of the equation (6) because the liquid-phase concentrations of the two components are similar. There is also a weak correlation between H_{Glu} and b_{Glu}, and between H_{Fru} and b_{Fru}. These trends have been confirmed in a previous study (Borg et al., 2013), which is because both parameters appear in the equation (6) and thus resolving the correlations to identify each parameter independently is difficult. These correlations may be weakened by increasing the number of data sets under different experimental conditions in the estimation.

A time-series graph of concentration is shown in Figure 4, superimposing the artificial data and the model which were calculated by substituting 5,000 values sampled randomly from the parameter posterior distributions. The simulation and experimental data fit well, suggesting that the estimation was done correctly.

4.1.1. Model validation by individual component deconvolution

We validated the model using the estimated parameters to confirm it could correctly deconvolute the overall concentration into individual concentrations of the two components, glucose and fructose. As described in Section 3.3.2 and Table 2, the estimation used two sets of single-component pulse injections, (A) and (B), in addition to the two-component step injection (C) which only gives the *overall* (total) concentration; we attempt to estimate the *individual* concentrations of the two components in (C) using the obtained model with the estimated parameters. The estimated individual concentrations are compared against the unused data sets of individual concentrations, which are assumed to be obtained using another measurement technique such as analytical HPLC.

Figure 4 (d) shows a time-series graph of the concentrations of glucose and fructose in (C), superimposed on the artificial data, and the deconvoluted concentrations given by the model using 5,000 randomly sampled values from the estimated parameter posterior distributions. It can be seen that the artificial data and deconvoluted concentrations given by the model overlap, indicating that the deconvolution is successful. This successful estimation is because the experimental data for parameter estimation was prepared well; that is, the information about H_{Glu}, H_{Fru}, K_{Glu}, and K_{Fru} can be obtained from single-component experiments, (A) and (B), respectively, and the information on b_{Glu} and b_{Fru} can be obtained from a high concentration experiment, (C) as reported in the previous study (Felinger et al., 2003). It can be seen in Equation (6) that b_{Glu} and b_{Fru} become sensitive only if the concentrations C_{Fru} and C_{Glu} and are sufficiently high; i.e. when $b_{Glu}C_{Glu}(x,t)$ >>0 and $b_{Fru}C_{Fru}(x,t)$>>0, which is realized in (C).

It is worth noting that this successful deconvolution of the online overall concentrations, which is enabled by the model and successful parameter estimation, may reduce the experimental effort substantially. In the current practice of chromatography modeling, experimentalists often fractionate the effluent from the column outlet and analyze the samples to *experimentally deconvolute* the concentrations. The number of samples tends to be large, which may take a substantial amount of time for HPLC analysis. This time-consuming practice is considered necessary to improve the accuracy of parameter estimation. Our case study, on the other hand, demonstrated that the overall concentrations obtained easily online by RI detectors can be *numerically deconvoluted* into individual concentrations.

4.1.2. Comparison between SMC and MCMC

We confirm that the estimation given by SMC is identical to that by MCMC. The modes and 95% credible intervals of the posterior distributions estimated by the two methods are shown in Table 3, along with the true values of the parameters. Furthermore, superimposed histograms showing the posterior distributions of each parameter obtained by SMC and MCMC are shown in Figure 5. All of

these results show that the estimations given by SMC and MCMC are identical. It can also be seen in Table 3 that the true values are within the 95% credible interval for all estimation results. Looking at the values in more detail, it can be seen that for Henry's constants H_i and overall mass transfer coefficients K_i, the modes are close to the true values, and the 95% credible interval is within 1% of the true value. On the other hand, for the adsorption equilibrium constants b_i, the modes slightly deviate from the true values, and the 95% credible interval is wider than 1% of the true value. A very similar trend was reported in a previous study (Grosfils et al., 2010), except for the overall mass transfer coefficient, which may be due to the different sensitivity of the parameters to the model. In our estimation, the slight deviations of the modes of parameter distributions from their true values may be due to the artificial random error added to the simulated artificial data of 300 time-series points. If the data generation and estimation had been repeated for a sufficiently large number of times, the mode of the parameter distributions would have matched the true values.

The computation time of SMC is significantly shorter than that of MCMC. The SMC estimation requires 34.9 minutes, which is less than one-tenth of that in the MCMC estimation, 352 minutes. Since the calculation times vary depending on the conditions of the simulation, it is not appropriate to simply compare these times. However, since it is easy to implement parallel computing in SMC, the computation time of SMC can be easily reduced using a many-core computer. The impact of the number of cores on the computation time of SMC is discussed in a later section.

Table 3. True values, modes, and 95% credible intervals of parameters for artificial data

Parameter	True value	Prior distributions	Estimated value by SMC	Estimated value by MCMC	Units
H_{Glu}	0.301	$0.301^{+2.96\times10^{-2}}_{-2.96\times10^{-2}}$	$0.301^{+9.07\times10^{-4}}_{-9.15\times10^{-4}}$	$0.301^{+7.47\times10^{-4}}_{-9.15\times10^{-4}}$	–
H_{Fru}	0.531	$0.532^{+5.21\times10^{-2}}_{-5.21\times10^{-2}}$	$0.532^{+1.24\times10^{-3}}_{-7.77\times10^{-4}}$	$0.532^{+8.03\times10^{-3}}_{-1.04\times10^{-3}}$	–
K_{Glu}	4.70×10^{-3}	$4.71\times10^{-3}{}^{+4.63\times10^{-4}}_{-4.63\times10^{-4}}$	$4.71\times10^{-3}{}^{+2.64\times10^{-5}}_{-3.41\times10^{-5}}$	$4.71\times10^{-3}{}^{+2.74\times10^{-5}}_{-2.84\times10^{-5}}$	s^{-1}
K_{Fru}	8.30×10^{-3}	$8.30\times10^{-3}{}^{+8.13\times10^{-4}}_{-8.13\times10^{-4}}$	$8.29\times10^{-3}{}^{+5.26\times10^{-5}}_{-4.63\times10^{-5}}$	$8.29\times10^{-3}{}^{+5.52\times10^{-5}}_{-3.89\times10^{-5}}$	s^{-1}
b_{Glu}	6.34×10^{-4}	$5.00\times10^{-4}{}^{+4.75\times10^{-4}}_{-4.75\times10^{-4}}$	$6.70\times10^{-4}{}^{+4.58\times10^{-5}}_{-8.77\times10^{-5}}$	$6.50\times10^{-4}{}^{+5.85\times10^{-5}}_{-6.31\times10^{-5}}$	L/g
b_{Fru}	2.48×10^{-4}	$5.00\times10^{-4}{}^{+4.75\times10^{-4}}_{-4.75\times10^{-4}}$	$2.20\times10^{-4}{}^{+7.61\times10^{-5}}_{-6.76\times10^{-5}}$	$2.27\times10^{-4}{}^{+5.60\times10^{-5}}_{-6.66\times10^{-5}}$	L/g
σ	5.00×10^{-3}	$4.67\times10^{-3}{}^{+9.15\times10^{-4}}_{-9.15\times10^{-4}}$	$4.72\times10^{-3}{}^{+4.17\times10^{-4}}_{-3.09\times10^{-4}}$	$4.69\times10^{-3}{}^{+4.00\times10^{-4}}_{-3.07\times10^{-4}}$	–

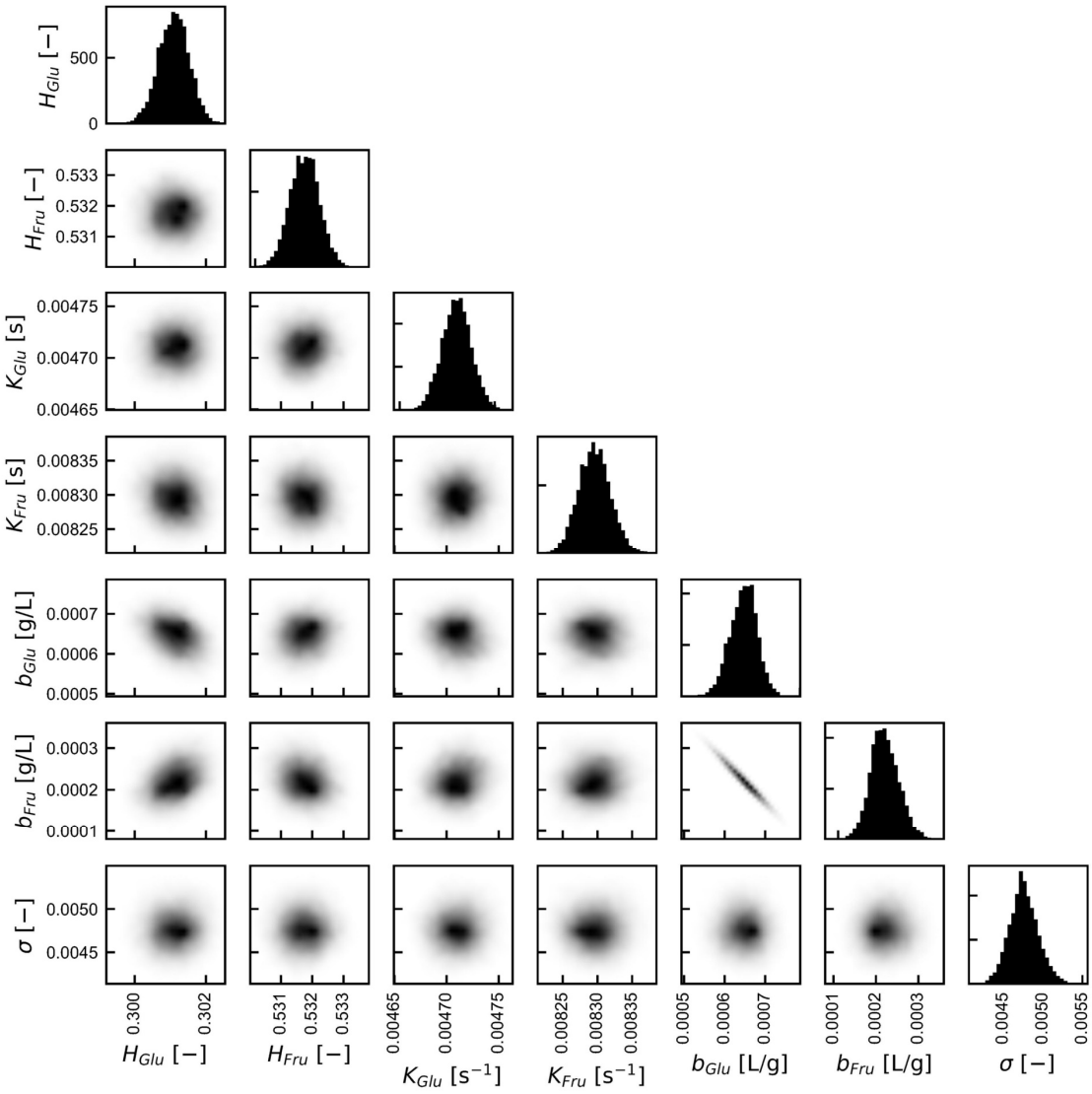

Figure 3. Density plot matrix of the posterior distributions of the parameters using the artificial data estimated by SMC. Each of the diagonal histograms shows the posterior distribution of each parameter, and each of the non-diagonal density plots shows the correlation between two parameters.

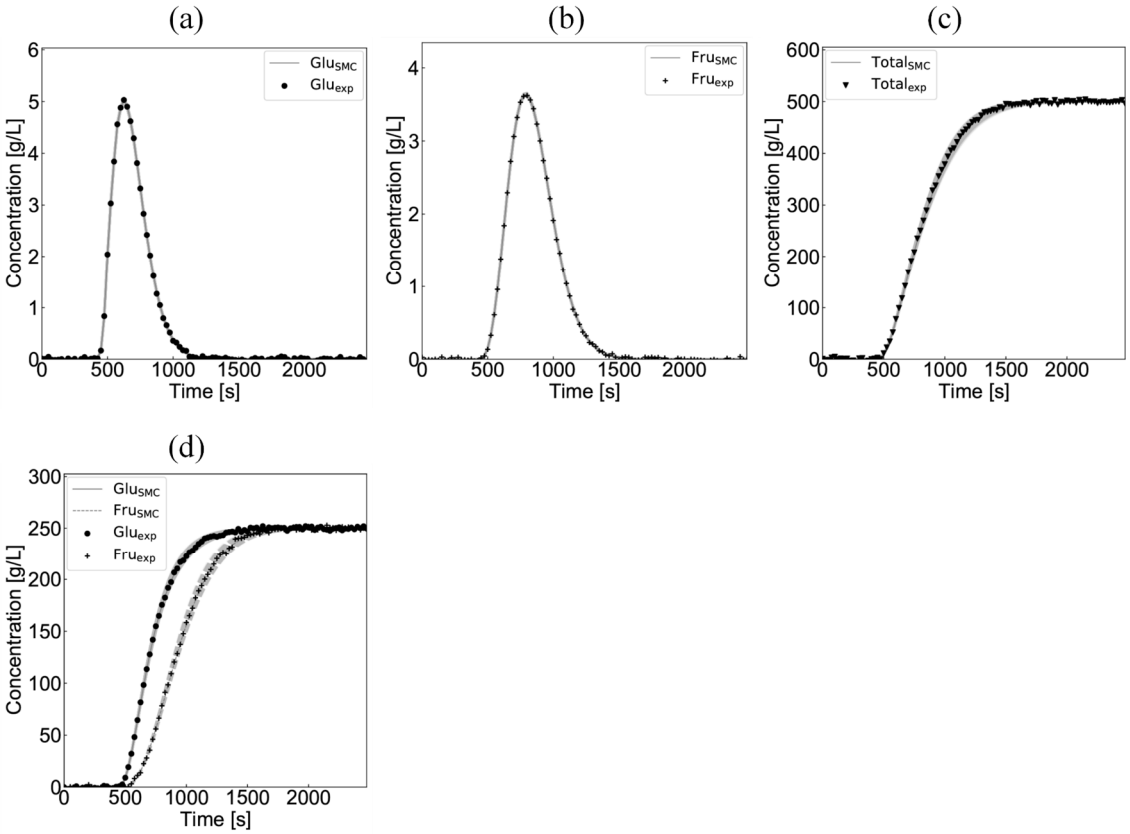

Figure 4. Fittings of simulations using the estimated parameters and the artificial data used for estimation((a)-(c)) and validation((d)). (a) glucose concentration obtained in experiment (A), (b) fructose concentration obtained in experiment (B), (c) total concentration of glucose and fructose obtained in experiment (C), and (d) validation test of parameter estimation results using deconvoluted concentration. The gray lines are given by the model using 5,000 randomly sampled parameter values from the estimated parameter posterior distributions."

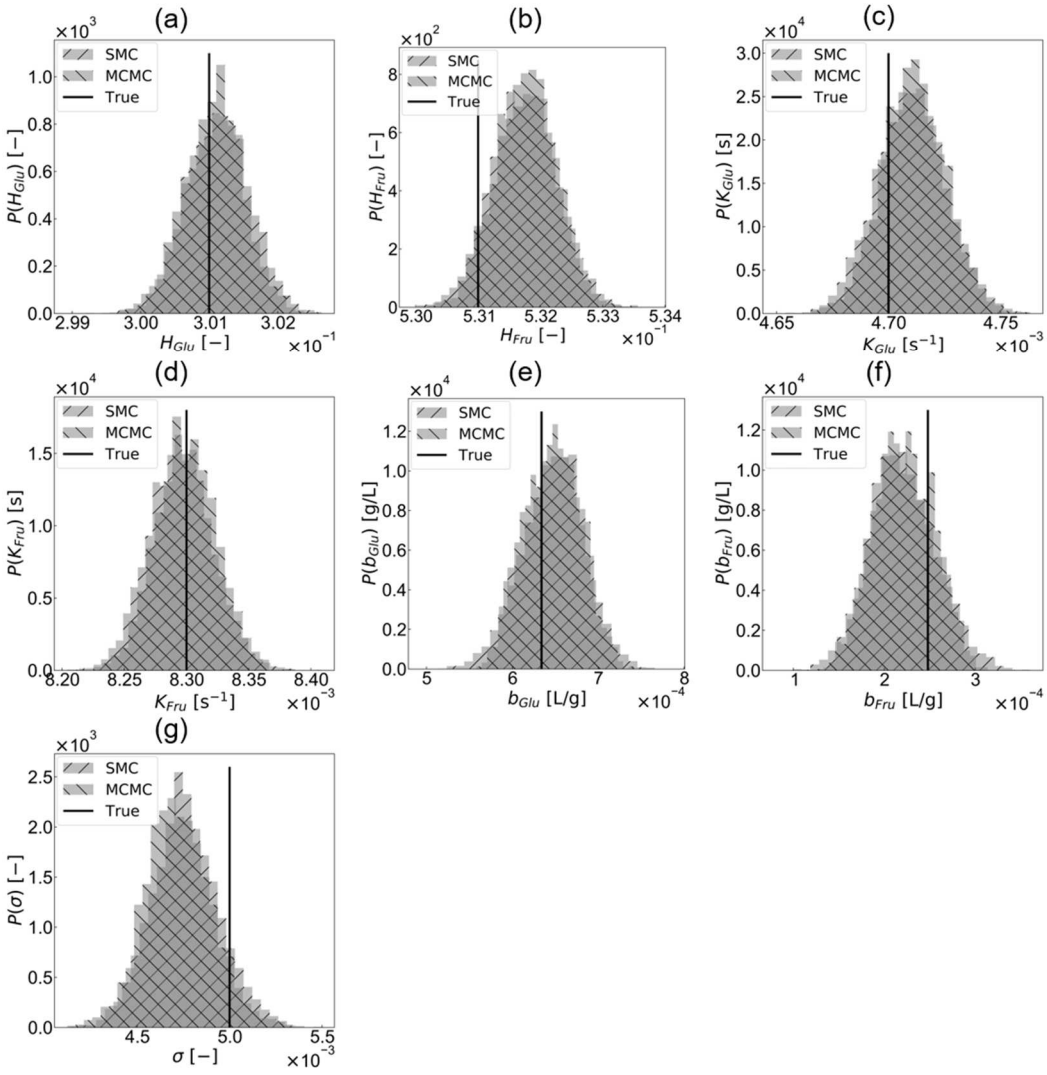

Figure 5. Histograms for the posterior distributions of each parameter obtained from SMC and MCMC using the artificial data. (a) H_{Glu}, (b) H_{Fru}, (c) K_{Glu}, (d) K_{Fru}, (e) b_{Glu}, (f) b_{Fru}, (g) σ.

4.2. Case.2: Lab-scale experimental data

The purpose of this case study is to verify the practical applicability of the proposed method by estimating the parameters and quantifying the uncertainty by SMC using actual lab-scale experimental data. The density plot matrix of the posterior distributions of the parameters obtained by SMC is shown in Figure 6. The modes and 95% credible intervals of the posterior distributions of the parameters estimated by SMC are shown in Table 4. The histograms and the shapes of the correlations among the parameters from Figure 6 are generally similar to those of the artificial data.

However, the value of b_{Fru} converges toward 0, the lower bound of the uniform prior distribution, and the mode of b_{Fru} is two orders of magnitude smaller than one of b_{Glu} (Table 4). This result may

not be justified considering the two molecules, both of which are monosaccharides, have similar molecular structures. Indeed, in a past study (Nowak et al., 2007), b_{Glu} and b_{Fru} are estimated to be within the same order of magnitude. The reason for the potential poor estimation of this parameter in our study may be that the sensitivity of the adsorption equilibrium constant is small in our experimental conditions. This is also supported by the 95% credible interval of b_{Fru} in the case study of the artificial data in Section 4.1, which was significantly wider than that of the other parameters. This potentially low sensitivity may be because there was only one type of breakthrough test conducted at the concentration of 250 g/L in this study (experiment (C)), while in the previous study (Nowak et al., 2007) breakthrough tests were conducted under many conditions at concentrations of up to as high as 600 g/L. The challenge of estimating b_{Fru} under low concentration was exacerbated by the larger measurement error in the lab-scale experimental data (estimated mean of σ is 1.11×10^{-2}), than that in the artificial data (estimated mean of σ is 4.72×10^{-3}).

A time-series graph of concentration is shown in Figure 7, superimposing the experimental data and the simulation results, which were calculated by substituting 5,000 randomly sampled values from the estimated parameter posterior distribution into the model. The fit is generally good, but in the high concentration region of (c), the model with the estimated parameters gives smaller values than the experimental data. This is most likely due to errors in the experimental operation of dilution of the HPLC analysis, which also explains the observation that the experimental data in the breakthrough curve exceeded the concentration of the feed, 500 g/L at around 1000 s and onward.

4.2.1. Model validation by individual component deconvolution

Validation of the model was performed by comparing numerical and experimental component deconvolutions under the same conditions as in the artificial data. As in the case of the artificial data, a time-series graph of concentration showing the simulation results superimposed on the experimental data is shown in Figure **7** (d). As can be seen in this figure, the numerical deconvolution nearly overlaps the experimental deconvolution, indicating the accuracy of the model. The slight deviation in the high concentration region at around 1000 s and onward which can also be seen in Figure 7 (c), is probably caused by the dilution error in HPLC analysis, as discussed in the previous section. This successful estimation shows the numerical deconvolution using the model with the estimated parameters can eliminate the need for experimental deconvolution, which may reduce experimental effort substantially.

4.2.2. Variation in computation time due to the difference in the number of cores

We analyzed the relationship between the number of cores used in parallel and the overall computation time for SMC. Figure 8 shows as the number of cores used in the calculations are increased, the overall computation time decreases. However, the computation time reaches a plateau

when the number of cores is approximately 10.

The limitation in decreasing the computation time may be explained by several reasons. First, there are calculation blocks that cannot be parallelized. Blocks in which each particle can be calculated independently—such as model simulation and likelihood evaluation—can be processed in parallel, but blocks in which particles interfere with each other—such as the calculation of the normalized weights and resampling—must be calculated in series. Second, communication overhead between cores to send each particle to each core increases as the number of cores increases. Finally, tasks in some cores must wait for other cores to finish before proceeding to the next task. The decision of choosing the number of cores should be made carefully considering the limitations in parallel implementation discussed above.

Table 4. Modes, and 95% credible intervals of parameters for lab-scale experimental data

Parameter	Prior distributions	Estimated value by SMC	Units
H_{Glu}	$0.342^{+3.35\times10^{-2}}_{-3.35\times10^{-2}}$	$0.342^{+1.12\times10^{-3}}_{-1.71\times10^{-3}}$	–
H_{Fru}	$0.552^{+5.41\times10^{-2}}_{-5.41\times10^{-2}}$	$0.552^{+1.57\times10^{-3}}_{-1.61\times10^{-3}}$	–
K_{Glu}	$2.23\times10^{-2}{}^{+2.20\times10^{-3}}_{-2.20\times10^{-3}}$	$2.22\times10^{-2}{}^{+5.16\times10^{-4}}_{-4.12\times10^{-4}}$	s^{-1}
K_{Fru}	$2.78\times10^{-2}{}^{+2.72\times10^{-3}}_{-2.72\times10^{-3}}$	$2.78\times10^{-2}{}^{+4.84\times10^{-4}}_{-4.90\times10^{-4}}$	s^{-1}
b_{Glu}	$5.00\times10^{-4}{}^{+4.75\times10^{-4}}_{-4.75\times10^{-4}}$	$4.40\times10^{-4}{}^{+2.92\times10^{-5}}_{-4.89\times10^{-5}}$	L/g
b_{Fru}	$5.00\times10^{-4}{}^{+4.75\times10^{-4}}_{-4.75\times10^{-4}}$	$2.30\times10^{-6}{}^{+5.49\times10^{-5}}_{-1.81\times10^{-6}}$	L/g
σ	$1.09\times10^{-2}{}^{+2.14\times10^{-3}}_{-2.14\times10^{-3}}$	$1.11\times10^{-2}{}^{+1.01\times10^{-3}}_{-9.21\times10^{-4}}$	–

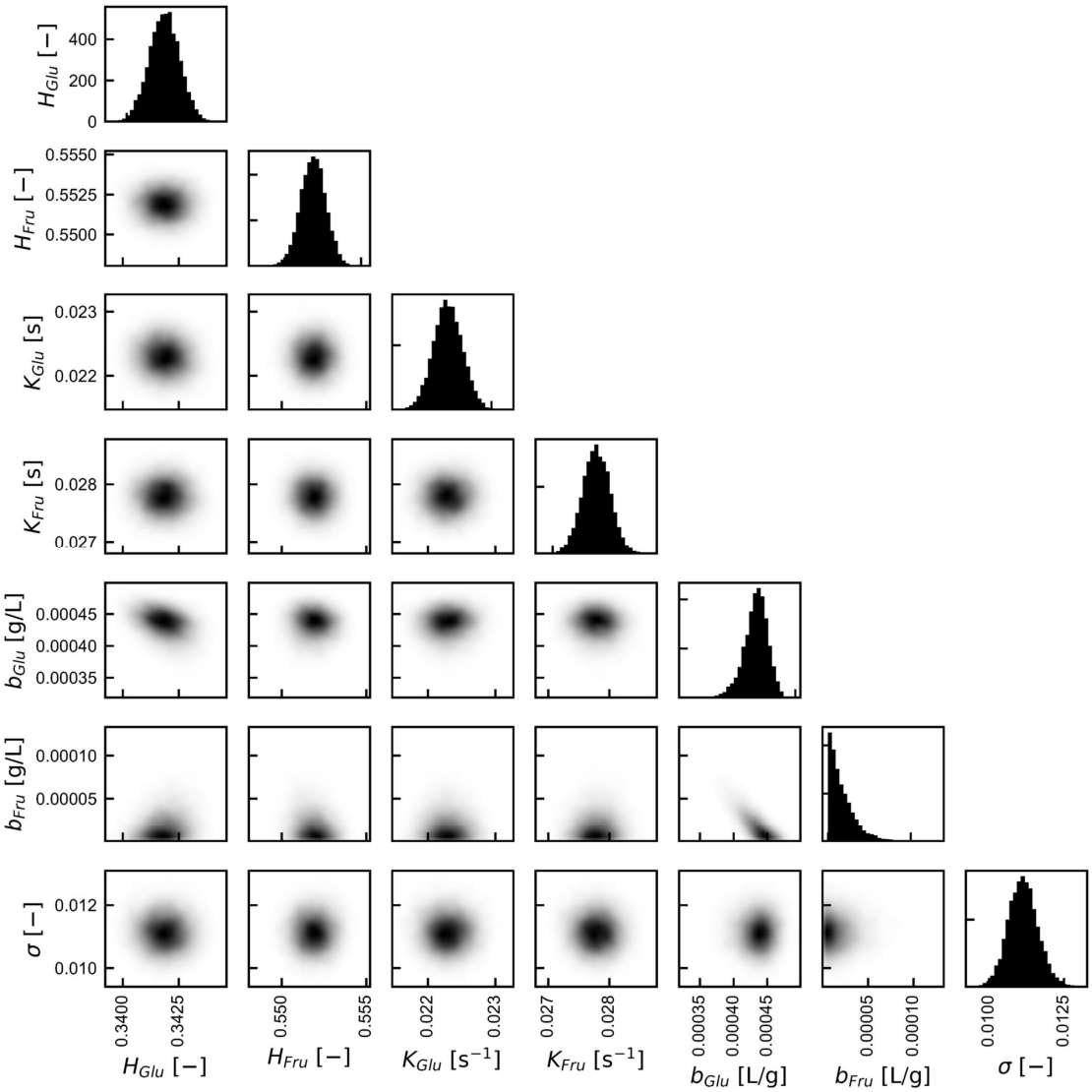

Figure 6. Density plot matrix of the posterior distributions of the parameters using the lab-scale experimental data estimated by SMC.

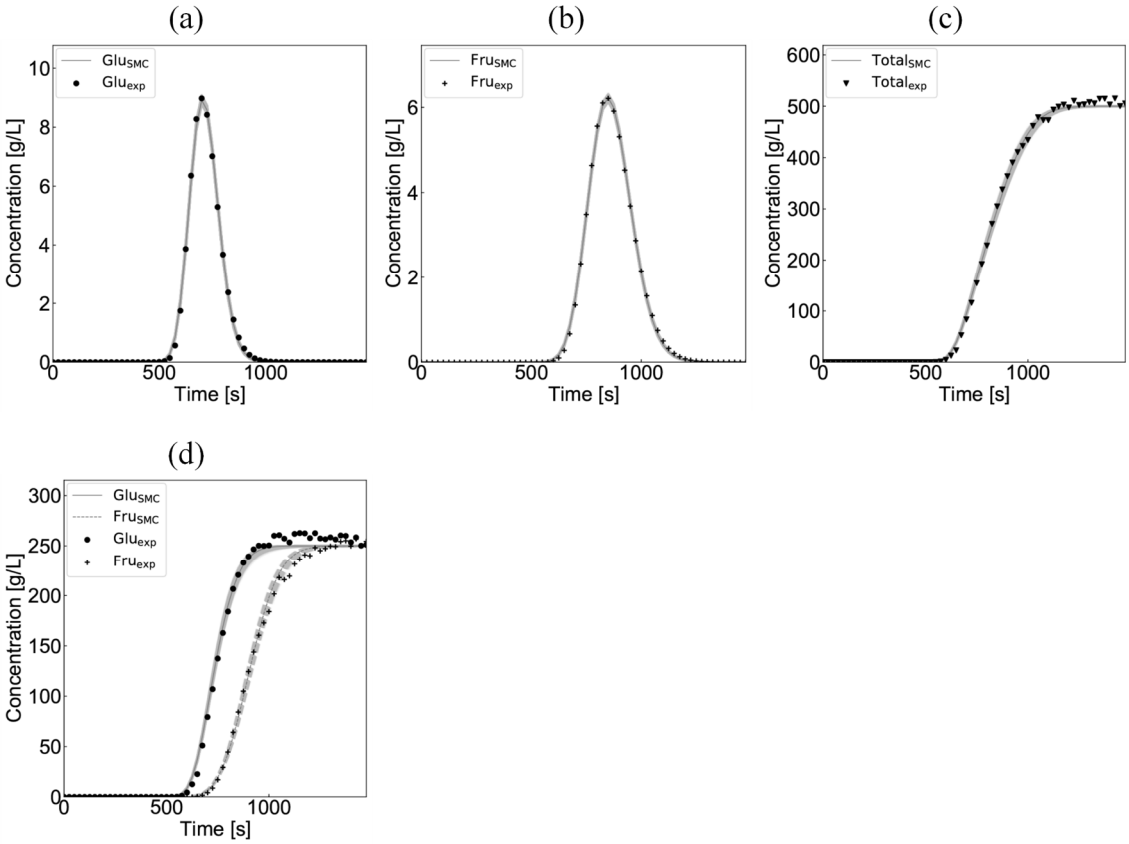

Figure 7. Fittings of simulations using the estimated parameters and the lab-scale experimental data used for estimation((a)-(c)) and validation((d)). (a) glucose concentration obtained in experiment (A), (b) fructose concentration obtained in experiment (B), (c) total concentration of glucose and fructose obtained in experiment (C), and (d) validation test of parameter estimation results using deconvoluted concentration obtained by HPLC analysis of fractionated samples (experimental deconvolution). The gray lines are given by the model using 5,000 randomly sampled parameter values from the estimated parameter posterior distributions."

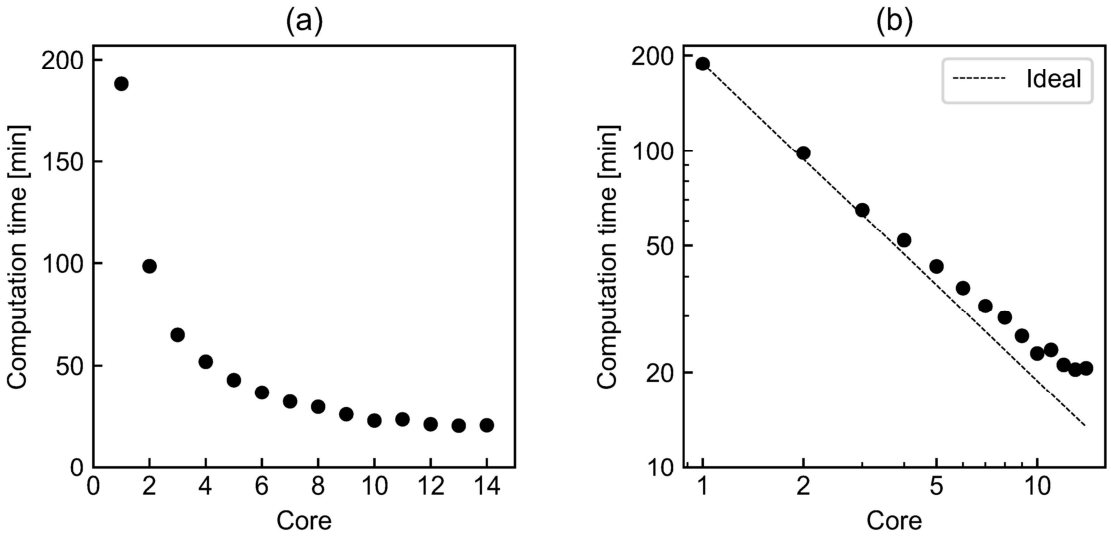

Figure 8. (a) Relationship between the number of cores and computation time, (b) log-log plot. The dotted line shows the ideal relationship, where the computation time is inversely proportional to the number of cores.

5. Conclusions

In this study, we developed an approach for UQ of parameters by SMC for chromatographic separation processes, and demonstrated using artificial data and lab-scale experimental data that the proposed approach can shorten the computation time. By applying the likelihood tempering SMC to the chromatography model, we confirmed that the estimation accuracy is nearly identical to that of MCMC while reducing the computation time substantially. Also, we demonstrated that the model with the estimated parameters can deconvolute the overall concentrations into individual concentrations, eliminating the need for extra experimental effort. Furthermore, we investigated the relationship between the number of cores used for parallel computation and identified the limitation in reducing the computation time by increasing the number of cores.

We have also identified some issues that need to be considered in future work. First, to improve the estimation accuracy by weakening the correlation between parameters, it is necessary to have multiple data sets under different conditions such as injected volume and feed concentration. Design of experiments (Franceschini and Macchietto, 2008; Kalyanaraman et al., 2014) may be utilized in determining the experimental conditions. Second, although the computation time has been considerably reduced, more challenging processes, such as SMB, exist. For this problem, surrogate models using neural networks (Anthony and Bartlett, 2009; Wu and Zhao, 2018) or support vector regression (Smola and Schölkopf, 2004; Song et al., 2002) can be potential solutions. Finally, further model validation tests should be performed for different process designs, such as the one with recycle

stream, and multi-column processes.

Acknowledgements

This study was supported by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research, Grant Number JP18H01776.

Notation

b_i	adsorption equilibrium constant for component i [L/g]
C_i	concentration of component i in the liquid phase [g/L]
$C_{i,feed}$	concentration of component i in the feed in the liquid phase
$C_{i,in}$	concentration of component i in the liquid phase flowing into the column
C_i^{eq}	concentration of component i in the liquid phase at the adsorption equilibrium [g/L]
$\boldsymbol{C}_j{}^{data}$	*j*th data of concentration in the liquid phase gotten by the experiments [g/L]
$\boldsymbol{C}_j{}^{model}$	*j*th data of concentration in the liquid phase calculated by the simulation model [g/L]
H_i	Henry's constant for component i [−]
K_i	overall mass transfer coefficient of component i [s^{-1}]
L	total steps of SMC
M	number of observations of the data
$\boldsymbol{N}$	Normal distribution
N_P	total number of particles
P	probability distribution
$\boldsymbol{Q}$	proposal distribution
q_i	concentration of component i in the solid phase [g/L]
q_i^{eq}	concentration of component i in the solid phase at the adsorption equilibrium [g/L]
t	time [s]
t_f	injection time of the feed [s]
u	superficial velocity of the liquid [g/L]
V_{inject}	total injected volume of feed
x	coordinate in the axial direction [m]
$\boldsymbol{y}$	observed data ($\boldsymbol{y} = \{(t_k, \boldsymbol{C}_k)\}_{k=1}^{M}$)
w_l^m	weight of the mth particle at lth step of SMC

$\widetilde{w}_l^m$	normalized weight of the mth particle at lth step of SMC

Greek letters

α	adoption rate
β	uniform random number generated as $\beta \sim U(0,1)$
γ_l	monotonically increasing scalar value of SMC
ε_b	overall porosity in the column $[-]$
$\boldsymbol{\theta}$	parameter to be estimated ($\boldsymbol{\theta} = [H_{Glu}, H_{Fru}, K_{Glu}, K_{Fru}, b_{Glu}, b_{Fru}, \sigma]^{\mathrm{T}}$)
$\boldsymbol{\theta}_k$	parameter adopted in the kth sampling of the MCMC sampling
$\boldsymbol{\theta}^{MAP}$	parameter calculated by maximum a posteriori estimation
$\boldsymbol{\theta}^{mle}$	parameter calculated by maximum likelihood estimation
$\boldsymbol{\theta}_0$	initial value of the parameter of the MCMC sampling
$\boldsymbol{\theta}^*$	candidate parameter value of the MCMC sampling
π_l	distribution of the parameters at the lth step of SMC
$\boldsymbol{\Sigma}_k$	covariance of the kth proposal distribution
σ	standard deviation assuming that the observation errors of all data follow a constant normal distribution $[-]$

Superscripts

$data$	experimental data
eq	equilibrium
m	particle number
MAP	maximum a posteriori estimation
mle	maximum likelihood estimation
$model$	simulation data computed by the LDF model
$*$	candidate value of the Metropolis sampling

subscripts

$feed$ or f	feed
Fru	fructose component
Glu	glucose component
i	component
in	inflow to the column
$inject$	injection
j	step of data
k	step of the MCMC sampling
l	step of SMC

References

Anthony, M., Bartlett, P.L., 2009. Neural Network Learning: Theoretical Foundations. Cambridge University Press, Cambridge.

Barbu, A., Zhu, S.-C., 2020. Monte Carlo Methods: Springer Nature. https://dx.doi.org/10.1007/978-981-13-2971-5

Bentley, J., Kawajiri, Y., 2013. Prediction-correction method for optimization of simulated moving bed chromatography. AIChE Journal 59, 736–746. https://doi.org/10.1002/aic.13856

Borg, N., Westerberg, K., Andersson, N., von Lieres, E., Nilsson, B., 2013. Effects of uncertainties in experimental conditions on the estimation of adsorption model parameters in preparative chromatography. Computers & Chemical Engineering 55, 148–157. https://doi.org/10.1016/j.compchemeng.2013.04.013

Briskot, T., Stückler, F., Wittkopp, F., Williams, C., Yang, J., Konrad, S., Doninger, K., Griesbach, J., Bennecke, M., Hepbildikler, S., Hubbuch, J., 2019. Prediction uncertainty assessment of chromatography models using Bayesian inference. Journal of Chromatography A 1587, 101–110. https://doi.org/10.1016/j.chroma.2018.11.076

Broughton, D.B., Gerhold, C.G., 1961. Continuous sorption process employing fixed bed of sorbent and moving inlets and outlets. US2985589A.

Chilamkurthi, S., Willemsen, J.-H., van der Wielen, L.A.M., Poiesz, E., Ottens, M., 2012. High-throughput determination of adsorption equilibria for chromatographic oligosaccharide separations. Journal of Chromatography A 1239, 22–34. https://doi.org/10.1016/j.chroma.2012.03.042

Conrad, P.R., Davis, A.D., Marzouk, Y.M., Pillai, N.S., Smith, A., 2018. Parallel Local Approximation MCMC for Expensive Models. SIAM/ASA J. Uncertainty Quantification 6, 339–373. https://doi.org/10.1137/16M1084080

Ettre, L.S., Sakodynskii, K.I., 1993. M. S. Tswett and the discovery of chromatography II: Completion of the development of chromatography (1903–1910). Chromatographia 35, 329–338. https://doi.org/10.1007/BF02277520

Felinger, A., Cavazzini, A., Guiochon, G., 2003. Numerical determination of the competitive isotherm of enantiomers. Journal of Chromatography A 986, 207–225. https://doi.org/10.1016/S0021-9673(02)01919-2

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. Chemical Engineering Science, Model-Based Experimental Analysis 63, 4846–4872. https://doi.org/10.1016/j.ces.2007.11.034

Francotte, E.R., Richert, P., 1997. Applications of simulated moving-bed chromatography to the separation of the enantiomers of chiral drugs. Journal of Chromatography A 769, 101–107. https://doi.org/10.1016/S0021-9673(97)00172-6

Froehlich, F., Theis, F., Hasenauer, J., 2014. Uncertainty Analysis for Non-identifiable Dynamical

Systems: Profile Likelihoods, Bootstrapping and More. https://doi.org/10.1007/978-3-319-12982-2_5

Gao, M., Zhang, H., 2012. Sequential Monte Carlo methods for parameter estimation in nonlinear state-space models. Computers & Geosciences 44, 70–77. https://doi.org/10.1016/j.cageo.2012.03.013

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian Data Analysis, third ed. Chapman and Hall/CRC, Boca Raton.

Gordon, N.J., Salmond, D.J., Smith, A.F.M., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings F - Radar and Signal Processing 140, 107–113. https://doi.org/10.1049/ip-f-2.1993.0015

Grosfils, V., Hanus, R., Wouwer, A.V., Kinnaert, M., 2010. Parametric uncertainties and influence of the dead volume representation in modelling simulated moving bed separation processes. Journal of Chromatography A 1217, 7359–7371. https://doi.org/10.1016/j.chroma.2010.09.030

Grosfils, V., Levrie, C., Kinnaert, M., Vande Wouwer, A., 2007. A systematic approach to SMB processes model identification from batch experiments. Chemical Engineering Science 62, 3894–3908. https://doi.org/10.1016/j.ces.2007.04.015

Guiochon, G., Felinger, A., Dean G. Shirazi, Anita M. Katti, 2006. Fundamentals of Preparative and Nonlinear Chromatography, second ed. Academic Press, Boston.

Hashimoto, K., Adachi, S., Noujima, H., Maruyama, H., 1983. Models for the Separation of Glucose/Fructose Mixture Using a Simulated Moving-Bed Adsorber. Journal of Chemical Engineering of Japan 16, 400–406. https://doi.org/10.1252/jcej.16.400

Hastings, W.K., 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. Biometrika 57, 97-109. https://doi.org/10.1093/biomet/57.1.97

He, Q.-L., Zhao, L., 2020. Bayesian inference based process design and uncertainty analysis of simulated moving bed chromatographic systems. Separation and Purification Technology 246, 116856. https://doi.org/10.1016/j.seppur.2020.116856

Horvath, C.G., Preiss, B.A., Lipsky, S.R., 1967. Fast liquid chromatography. Investigation of operating parameters and the separation of nucleotides on pellicular ion exchangers. Anal. Chem. 39, 1422–1428. https://doi.org/10.1021/ac60256a003

Hsu, S.-H., Stamatis, S. D., Caruthers, J. M., Delgass, W. N., Venkatasubramanian, V., Blau, G. E., Lasinski, M., Orcun, S., 2009. Bayesian framework for building kinetic models of catalytic systems. Industrial & Engineering Chemistry Research, 48, 4768-4790. https://doi.org/10.1021/ie801651y

Joshi, M., Seidel-Morgenstern, A., Kremling, A., 2006. Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. Metabolic Engineering 8,

447–455. https://doi.org/10.1016/j.ymben.2006.04.003

Kaltenbrunner, O., Giaverini, O., Woehle, D., Asenjo, J.A., 2007. Application of chromatographic theory for process characterization towards validation of an ion-exchange operation. Biotechnology and Bioengineering 98, 201–210. https://doi.org/10.1002/bit.21358

Kalyanaraman, J., Kawajiri, Y., Lively, R.P., Realff, M.J., 2016. Uncertainty quantification via bayesian inference using sequential monte carlo methods for CO2 adsorption process. AIChE Journal 62, 3352–3368. https://doi.org/10.1002/aic.15381

Kalyanaraman, J., Kawajiri, Y., Realff, M.J., 2014. Bayesian Estimation, Uncertainty Propagation and Design of Experiments for CO2 Adsorption on Amine Sorbents, in: Eden, M.R., Siirola, J.D., Towler, G.P. (Eds.), Computer Aided Chemical Engineering, Proceedings of the 8 International Conference on Foundations of Computer-Aided Process Design. Elsevier, pp. 345–350. https://doi.org/10.1016/B978-0-444-63433-7.50042-0

Kantas, N., Doucet, A., Singh, S.S., Maciejowski, J.M., 2009. An Overview of Sequential Monte Carlo Methods for Parameter Estimation in General State-Space Models. IFAC Proceedings Volumes, 15th IFAC Symposium on System Identification 42, 774–785. https://doi.org/10.3182/20090706-3-FR-2004.00129

Kawajiri, Y., 2021. Model-based optimization strategies for chromatographic processes: a review. Adsorption 27, 1–26. https://doi.org/10.1007/s10450-020-00251-2

Kawajiri, Y., Biegler, L.T., 2006. Optimization strategies for simulated moving bed and PowerFeed processes. AIChE Journal 52, 1343–1350. https://doi.org/10.1002/aic.10736

Kitagawa, G., 1996. Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models. Journal of Computational and Graphical Statistics 5, 1–25. https://doi.org/10.1080/10618600.1996.10474692

Lam, S.K., Pitrou, A., Seibert, S., 2015. Numba: a LLVM-based Python JIT compiler, in: Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, LLVM '15. Association for Computing Machinery, New York, NY, USA, pp. 1–6. https://doi.org/10.1145/2833157.2833162

Lang, L., Chen, W., Bakshi, B.R., Goel, P.K., Ungarala, S., 2007. Bayesian estimation via sequential Monte Carlo sampling—Constrained dynamic systems. Automatica 43, 1615–1622. https://doi.org/10.1016/j.automatica.2007.02.012

Leeuwen, P.J. van, 2009. Particle Filtering in Geophysical Systems. Monthly Weather Review 137, 4089–4114. https://doi.org/10.1175/2009MWR2835.1

Maeder, Marcel., Zuberbuehler, A.D., 1990. Nonlinear least-squares fitting of multivariate absorption data. Anal. Chem. 62, 2220–2224. https://doi.org/10.1021/ac00219a013

Martin, A.J.P., Synge, R.L.M., 1941. A new form of chromatogram employing two liquid phases: A theory of chromatography. 2. Application to the micro-determination of the higher

monoamino-acids in proteins. Biochemical Journal 35, 1358–1368. https://doi.org/10.1042/bj0351358

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of State Calculations by Fast Computing Machines. J. Chem. Phys. 21, 1087–1092. https://doi.org/10.1063/1.1699114

Moon, H., Kook Lee, W., 1986. A lumped model for multicomponent adsorptions in fixed beds. Chemical Engineering Science 41, 1995–2004. https://doi.org/10.1016/0009-2509(86)87116-0

Na, J., Park, S., Bak, J. H., Kim, M., Lee, D., Yoo, Y., Kim, I., Park, J., Lee, U., Lee, J. M., 2019. Bayesian Inference of Aqueous Mineral Carbonation Kinetics for Carbon Capture and Utilization. Industrial & Engineering Chemistry Research, 58, 8246-8259. https://doi.org/10.1021/acs.iecr.9b01062

Nowak, J., Gedicke, K., Antos, D., Piątkowski, W., Seidel-Morgenstern, A., 2007. Synergistic effects in competitive adsorption of carbohydrates on an ion-exchange resin. Journal of Chromatography A 1164, 224–234. https://doi.org/10.1016/j.chroma.2007.07.022

Rawlings, J. B., Bakshi, B. R., 2006. Particle filtering and moving horizon estimation. Computers & Chemical Engineering, 30, 1529-1541. https://doi.org/10.1016/j.compchemeng.2006.05.031

Schiesser, W.E., 1991. The Numerical Method of Lines: Integration of Partial Differential Equations. Academic Press, San Diego.

Schmidt-Traub, H., Schulte, M., Seidel-Morgenstern, A. (Eds.), 2012. Preparative Chromatography, second ed. Wiley-VCH.

Sherlock, C., Golightly, A., Henderson, D.A., 2017. Adaptive, Delayed-Acceptance MCMC for Targets With Expensive Likelihoods. Journal of Computational and Graphical Statistics 26, 434–444. https://doi.org/10.1080/10618600.2016.1231064

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Statistics and Computing 14, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Song, M., Breneman, C.M., Bi, J., Sukumar, N., Bennett, K.P., Cramer, S., Tugcu, N., 2002. Prediction of Protein Retention Times in Anion-Exchange Chromatography Systems Using Support Vector Regression. J. Chem. Inf. Comput. Sci. 42, 1347–1357. https://doi.org/10.1021/ci025580t

Sreedhar, B., Kawajiri, Y., 2014. Multi-column chromatographic process development using simulated moving bed superstructure and simultaneous optimization – Model correction framework. Chemical Engineering Science 116, 428–441. https://doi.org/10.1016/j.ces.2014.05.004

Uppsala University, 2017. Sequential Monte Carlo methods [WWW Document]. URL http://www.it.uu.se/research/systems_and_control/education/2017/smc/schedule (accessed 1.29.21).

Wang, S.-C., Tien, C., 1982. Further work on multicomponent liquid phase adsorption in fixed beds. AIChE Journal 28, 565–573. https://doi.org/10.1002/aic.690280407

Westerberg, K., Hansen, E.B., Degerman, M., Hansen, T.B., Nilsson, B., 2012. Model-Based Process Challenge of an Industrial Ion-Exchange Chromatography Step. Chemical Engineering & Technology 35, 183–190. https://doi.org/10.1002/ceat.201000560

Wu, H., Zhao, J., 2018. Deep convolutional neural network model based chemical process fault diagnosis. Computers & Chemical Engineering 115, 185–197. https://doi.org/10.1016/j.compchemeng.2018.04.009

Yamamoto, Y., 2021. YamamotoYota/SMC_for_chromatographic_model [WWW Document]. URL https://github.com/YamamotoYota/SMC_for_chromatographic_model (accessed 2.2.21).

Zhang, Z., 1997. Parameter estimation techniques: a tutorial with application to conic fitting. Image and Vision Computing 15, 59–76. https://doi.org/10.1016/S0262-8856(96)01112-2