

기술노트

베이지안 보정 기법을 활용한 생물-물리-화학적 반응 동역학 모델 최적화:
미생물 성장-사멸과 응집 동역학에 대한 사례 연구

이병준^{1,2,†}

¹경북대학교 에너지환경연구소 · ²경북대학교 미래과학기술융합학과

Application of Bayesian Calibration for Optimizing Biophysicochemical Reaction Kinetics Models in Water Environments and Treatment Systems: Case Studies in the Microbial Growth-decay and Flocculation Processes

Byung Joon Lee^{1,2,†}

¹*Energy and Environment Institute, Kyungpook National University* ·

²*Department of Advanced Science and Technology Convergence, Kyungpook National University*

(Received 14 March 2024, Revised 2 June 2024, Accepted 4 June 2024)

Abstract

Biophysicochemical processes in water environments and treatment systems have been great concerns of engineers and scientists for controlling the fate and transport of contaminants. These processes are practically formulated as mathematical models written in coupled differential equations. However, because these process-based mathematical models consist of a large number of model parameters, they are complicated in analytical or numerical computation. Users need to perform substantial trials and errors to achieve the best-fit simulation to measurements, relying on arbitrary selection of fitting parameters. Therefore, this study adopted a Bayesian calibration method to estimate best-fit model parameters in a systematic way and evaluated the applicability of the calibration method to biophysicochemical processes of water environments and treatment systems. The Bayesian calibration method was applied to the microbial growth-decay kinetics and flocculation kinetics, of which experimental data were obtained with batch kinetic experiments. The Bayesian calibration method was proven to be a reasonable, effective way for best-fit parameter estimation, demonstrating not only high-quality fitness, but also sensitivity of each parameter and correlation between different parameters. This state-of-the-art method will eventually help scientists and engineers to use complex process-based mathematical models consisting of various biophysicochemical processes.

Key words : Bayesian calibration, Biophysicochemical process, Kinetics, Mathematical model, Parameter estimation

^{1,2} Corresponding author, 교수(Professor), bjlee@knu.ac.kr, <https://orcid.org/0000-0002-8338-5238>

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

물환경 및 수처리공정에서 발생하는 다양한 현상들을 이해하고 제어하기 위해서, 과학자 및 공학자들은 생물-물리-화학 공정(반응)들의 동역학을 규명하고 다양한 예측 모델을 수립하는 노력을 기울여왔다. 예를 들어, 하폐수처리 활성슬러지 공정에서 일어나는 생화학적 반응들을 이해하고 예측하기 위하여, 미국-유럽지역 연구자들을 중심으로 1980년대부터 활성슬러지모델(Activated Sludge Model, ASM)이란 연립미분방정식 기반 수학적 모델을 개발하여 왔다. 현재 ASM 모델은 하수처리장 공정 예측을 위한 대부분의 상용(Commercial) 소프트웨어의 기초 모델로 활용되고 있다(Duarte et al., 2024; Jeppsson, 1996; Van Loosdrecht et al., 2015). 다른 예로써, 물환경 및 수처리공정에서 발생하는 입자 응집 현상에 대하여, 입자 개체군수지식(Population Balance Equation, PBE)을 활용하여 물환경 부유입자 응집 반응을 연립미분방정식 형태의 수학적 모델로 구현하여 예측하고 있다(Kim and Lee, 2020; Lee et al., 2011). 이와 같이, 물환경 및 수처리공정 예측을 위한 다양한 노력이 이뤄져 왔으며, 그 결과 다양한 수학적 예측 모델이 개발, 활용되고 있다.

최근 동역학 기반 수학적 예측 모델과 함께, 물환경 및 수처리공정 예측을 위하여 머신러닝을 비롯한 데이터 기반 인공지능 기법이 활용되고 있다(Duarte et al., 2024). 인공지능 기법은 일종의 블랙박스 모델로써 생물-물리-화학적 공정(반응) 동역학에 기반하지 않고, 다양한 경로로 획득된 데이터를 기반으로 미래를 예측하는 모델이다. 즉, 시스템 내부에서 일어나는 매커니즘을 이해하지 않고서도, 충분한 데이터만 확보된다면 시스템 예측 결과를 도출할 수 있다. 하지만, 역설적으로, 데이터 기반 인공지능 기법은 데이터의 질적, 양적 기준을 확보하지 못하면 양질의 예측 결과를 도출할 수 없다. 예를 들어, 하수처리장 수질인자의 안정적인 관측 기술이 존재하지 않는다면, 하수처리장 공정 예측을 위한 데이터 기반 인공지능 기법의 활용에 제한이 있을 수 있다. 하지만, 생물-물리-화학 공정(반응) 동역학에 기반한 수학적 예측 모델은, 모델 검증 외 관측 데이터 의존성이 크지 않으며, 가상현실(디지털 트윈)로써 데이터와 무관하게 독립적으로 구동될 수 있다.

동역학 기반 수학적 모델은 다수의 데이터-모델 접합(Fitting) 매개변수(상수)를 가지며 이들 매개변수가 모델 예측 결과를 제어하므로, 사용자는 매개변수를 합리적으로, 엄격하게 산정하여 모델에 반영하여야 한다. 하지만, 다수 매개변수를 다루기는 쉽지 않은 작업이며, 대부분 사용자는 경험을 바탕으로 임의적으로 모델 매개변수를 선택한다. 하수처리장 활성슬러지공정 시뮬레이션을 위한 ASM 모델은 30~40여개의 매개변수를 가지며, 사용자가 다수 매개변수를 합리적으로 선택하기가 쉽지 않다(Gujer et al., 1999; Jeppsson, 1996). 최근 다양한 매개변수 평가 기법들이 개발되었으나, 본 연구에서는 최적 접합 매개변수 평가뿐만 아니라, 매개변수 간 상관관계 및 매개변수 변화에 따른 모델 민

감도(Sensitivity)까지 평가 가능한 베이저안 보정(Bayesian Calibration) 기법을 선택하였다(Vrugt, 2016; van Turnhout et al., 2016).

베이저안 보정 기법은 Bayes 정리를 기반한 매개변수 추정의 통계적 접근법으로써, 실험 데이터와 모델 예측치 간 최적 접합 매개변수를 산출한다. 베이저안 보정은 실험 데이터와 함께 매개변수의 사전 분포(Prior Distribution)를 기반으로 우도함수(Likelihood Function)를 계산함으로써 매개변수의 불확실성을 고려한 매개변수의 사후분포(Posterior Distribution)를 계산한다(Vrugt, 2016). 따라서, 베이저안 보정 기법은 실험-예측치 간 최적 접합 기준을 만족하는 다수의 매개변수 추정치들을 산출하며, 일반적으로 매개변수의 사후분포는 최빈값과 불확실성을 내포한 정규분포 형태가 된다. 베이저안 보정 기법은 매개변수의 불확실성이 큰 수문 모델(Hydrological Model)에 주로 활용되어 왔으며(Vrugt, 2016), 또한 생물/생태 동역학 모델에 적용된 바 있다(van Mourik et al., 2014; van Turnhout, 2016). 수문 및 생물/생태 동역학 모델은 복잡다단한 지구환경과 생체 현상을 다루므로, 모델 매개변수의 불확실성은 필연적으로 나타난다. 물환경 및 수처리 공정 또한 불확실성을 크게 내포하며, 베이저안 보정 기법을 활용한 매개변수 추정이 필요한 분야로 판단된다. 특히, 베이저안 보정 기법을 국내의 구축된 하천·호수 수질 모델(QUAL, WASP 등), 유역 환경 모델(HSPF, SWAT 등), 수처리 공정 모델(GPS-X, BIOWIN 등)과 연계 활용한다면, 불확실성을 내포한 모델 매개변수를 체계적, 과학적으로 평가할 수 있다.

본 연구는 베이저안 보정 기법을 물환경 및 수처리공정에서 일어나는 생물-물리-화학 공정(반응)의 대표 사례에 적용함으로써, 베이저안 보정 기법의 물환경 분야에서의 적용성을 평가, 소개하고자 한다. 하폐수처리장 활성슬러지공정의 종속영양미생물 성장-사멸 동역학과 물환경 부유입자 응집 동역학 배치 실험 데이터와 연립미분방정식 기반 수학적 모델을 활용하여, 최적 접합 매개변수 산정 방법으로써 베이저안 보정 기법의 물환경 분야 적용 가능성을 평가하였다. 물환경 분야의 연구자/공학자가 본 연구에 적용, 평가된 베이저안 보정 기법을 활용하게 되면, 기존 경험적, 임의적 매개변수 선택의 한계를 벗어나 이론적, 합리적인 매개변수 선택과 모델 운용을 할 수 있으리라 기대한다.

2. Materials and Methods

2.1 활성슬러지 종속영양미생물 성장-사멸 동역학 배치 실험

하수에 포함된 Readily Biodegradable COD (RBCOD)와 하수처리 시스템 활성슬러지의 호기성 미생물 분율을 정량 분석을 위하여 미생물 성장-사멸 동역학 실험(Microbial Growth Kinetic Test)을 수행하였다(Choi et al., 2017; Lee et al., 2006; Wentzel et al., 1995). Modified Ludzack-Ettinger (MLE) 타입 하수처리 시스템에서 추출한 활성슬러지와 하수를 혼합하여 3 L 회분식(Batch) 반응기에 투여 후 실험을

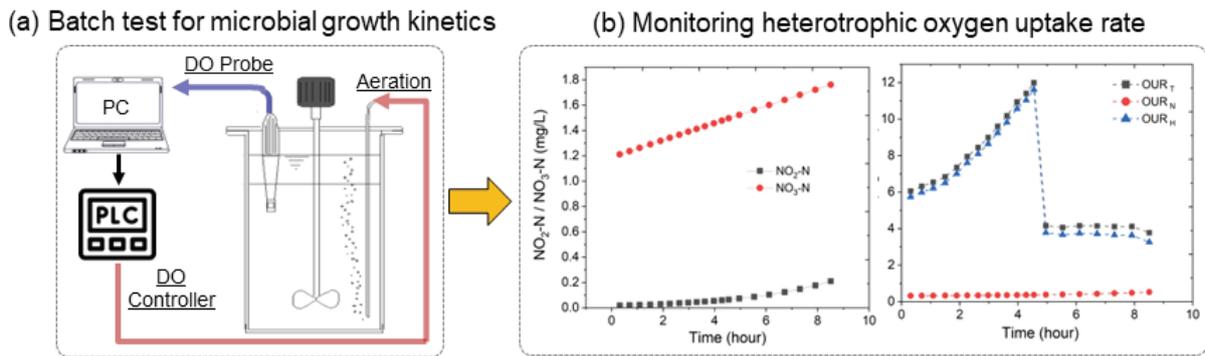


Fig. 1. (a) Schematic diagram of a batch test for microbial growth kinetics, and (b) Experimental results obtained from the batch test (i.e., changes of nitrate and nitrite concentration and oxygen uptake rate). OUR_T indicates total oxygen uptake rate. OUR_N and OUR_H indicate oxygen uptake rates by nitrification and heterotrophic microorganisms, respectively.

진행하였다(Fig. 1a). 반응기 용존산소 측정을 위한 용존산소 프로브(YSI Inc., 미국)와 산소 공급을 제어할 수 있는 PLC 시스템을 활용하여, 회분식 실험이 진행되는 동안 산소호흡율(Oxygen Uptake Rate, OUR)를 지속적으로 측정하였다. 용존산소 농도가 2 mgO/L 이하로 하락하면 산소공급이 자동으로 시작되며, 4 mgO/L 이상에 도달하면 산소공급이 차단되고, 산소공급이 중단된 기간 동안 측정된 용존산소 농도 저감율을 바탕으로 산소호흡율을 산정하였다. 또한, 회분식 실험이 진행되는 동안, 일정 시간 간격으로 여과 샘플을 확보하여 질산염 및 아질산염 분석을 수행하였고 질산화에 따른 산소호흡율을 산정하였다. 본 연구는 유기물 분해 종속영양미생물(Ordinary Heterotrophic Organisms)에 한정하여 미생물 성장-사멸 동역학 모델을 수립하였으므로, 질산화에 따른 산소호흡율(OUR_N)을 총 산소호흡율(OUR_T)에서 감하여 계산한 종속영양미생물에 의한 산소호흡율(OUR_H)을 모델 매개변수 보정 작업의 실험데이터로 활용하였다(Fig. 1). 회분식 실험 시작과 종료 시 샘플에 대한 COD 농도를 측정하여 COD 물질수지 수립에 활용하였다. 미생물 성장 동역학 실험의 산소호흡율 자료에 대하여 해석적 계산법(Analytical Solution)을 적용하면, 하수의 RBCOD 분율과 황성슬러지에

포함된 호기성 미생물 분율을 산출할 수 있다(Lee et al., 2014; Wentzel et al., 1995). 하지만, 본 연구는 해석적 계산법을 적용하지 않고 연립 미분방정식 형태의 수학적 모델을 베이지안 보정 기법과 연계하여 하수 RBCOD 분율과 황성슬러지 호기성 미생물 분율을 산정하였다.

2.2 응집 동역학 실험

응집 동역학의 최적 모델 계수 산정 도구로써 베이지안 보정 기법의 활용성을 평가하기 위하여 응집 동역학 테스트를 수행하였다. 생체고분자물질(Extracellular Polymeric Substances)의 일종인 Xanthan Gum (Sigma-Aldrich, USA)과 점토질 입자를 포함한 현탁액을 사용하여 2 L 표준 자테스터(Phipps and Bird, USA)에서 응집 동역학 테스트를 수행하였다(Kim and Lee, 2020). Xanthan Gum은 다당류 중합체로써 입자 응집을 증진시키는 것으로 알려져있다(Villacorte et al., 2015). 또한, 0.1~4.0 μm 크기 분포의 Kaolinite (Sigma-Aldrich, USA)를 점토질 입자로써 실험에 사용하였다. 응집 동역학 실험 시작 후 2분 동안, 반응조 현탁액을 200 rpm으로 급속 혼합하였으며, 이후 응집 동역학이 평형 상태에 도달할 때까지 현탁액의 난류 강도를 제어하기 위해 혼합속도를 100

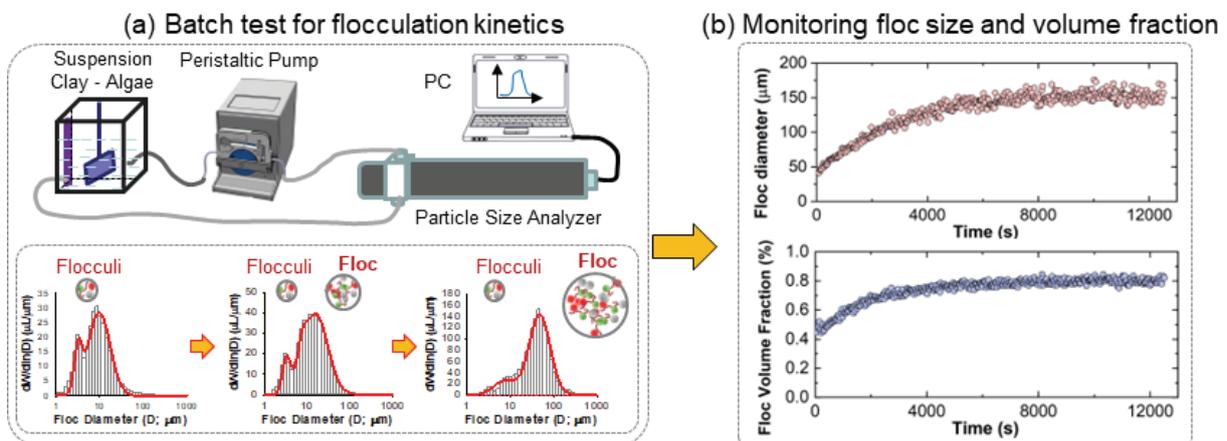


Fig. 2. (a) Schematic diagram of the batch test for flocculation kinetics and real-time measurement of floc size distribution, (b) Experimental results obtained from the batch test (i.e., mean diameter and volumetric fraction of flocs).

rpm으로 고정하여 교반하였다. 응집 동역학 실험을 진행하는 동안, 반응조와 광산란식 입도분석기(LISST-200X, Sequoia Inc., USA)를 6mm 튜브와 정량펌프로 연결하여 현탁액의 입도 분포를 실시간 관측하였다(Fig. 2a). LISST-200X 입도 분석기는 Mie 산란 이론에 기반하며 레이저가 수중 입자에 부딪힐 때 산란되는 빛을 분석하여 입도분포를 실시간 분석하는 장비로써, 본 연구의 연속류 흐름(Flow-through) 장치와 연계하여 매 30초 간격으로 반응조 내 입도분포를 관측하였다. 현탁액의 입자 농도가 높을 경우, 광산란식 입도분석기는 레이저 경로상 증첩에 의한 오류가 발생하므로, 본 연구의 점토질 현탁액의 농도는 0.12 g/L으로 제한하였다. 관측된 입도 분포는 응집 동역학 모델의 주요 종속변수인 응집체 평균 직경(D_f)과 부피 분율(V_f)을 환산되어, 응집 동역학 모델과 베이지안 보정 기법 적용성 평가에 활용되었다(Fig. 2b).

2.3 활성슬러지 종속영양미생물 성장-사멸 동역학 모델

본 연구의 생물학적 공정(반응) 동역학에 대한 베이지안 보정 기법 적용성 평가를 위하여, ASM1 (Activated Sludge Model No.1) 모델에 기반한 호기성 종속영양미생물 성장 동역학 모델을 활용하였다(Henze et al., 1987; Jeppsson, 1996). ASM1 호기성 미생물 성장 동역학 모델은 Monod 타입 모델을 기초로 최대 성장속도 상수와 반포화 상수를 포함한 성장 동역학과 1차 반응식 형태의 미생물 사멸(자가분해) 동역학으로 구성되며(Equation 1), 미생물 기질(먹이)은 빠르게 분해되는 Readily Biodegradable COD (RBCOD)와 느리게 분해되는 Slowly Biodegradable COD (SBCOD)로 구성된다(Equation 2~3). 특히, SBCOD는 미생물 가수분해(Hydrolysis) 작용에 의해 일정 분량은 RBCOD로 변환되며 나머지는 난분해성물질로 변환되는 것으로 가정된다. 미생물 성장 동역

학 모델은 각 호기성 미생물, RBCOD, SBCOD에 대한 연립 미분방정식 형태로 구성되며, 산소소모율(Oxygen Uptake Rate; OUR)은 시스템의 상태 변수로써 산출될 수 있다(Equation 4). 호기성 미생물 성장 동역학의 연립미분방정식은 Matlab (The MathWorks Inc., USA) 내장함수를 활용하여 적분되었으며, 베이지안 보정 기법과 연계하여 다수 모델 매개변수 조합에 대하여 연립미분방정식을 반복적으로 연산하였다. 호기성 미생물 성장 동역학의 연립미분방정식 계산을 위한 Matlab 코드는 Appendix 1~3에 수록하였다.

$$\frac{dX_H}{dt} = \mu_H \frac{S_S}{K_S + S_S} X_H - b_H X_H \quad (1)$$

$$\frac{dS_S}{dt} = -\frac{1}{Y_H} \mu_H \frac{S_S}{K_S + S_S} X_H + k_H \frac{(X_S/X_H)}{K_X + (X_S/X_H)} X_H \quad (2)$$

$$\frac{dX_S}{dt} = (1 - f_p) b_H X_H - k_H \frac{(X_S/X_H)}{K_X + (X_S/X_H)} X_H \quad (3)$$

$$\frac{dO}{dt} = OUR = -\frac{1 - Y_H}{Y_H} \mu_H \frac{S_S}{K_S + S_S} X_H \quad (4)$$

여기서, X_H = 호기성 미생물 농도 (mgCOD/L), S_S = Readily Biodegradable COD (RBCOD) 농도 (mgCOD/L), X_S = Slowly Biodegradable COD (SBCOD) 농도, Y_H = 기질의 생체 전환율 (-), μ_H = 최대 성장속도 상수 (/d), K_S = RBCOD 분해 반포화 상수 (mgCOD/L), k_H = 가수분해 속도 상수 (/d), K_X = SBCOD 분해 반포화 상수 (mgCOD/L), f_p = SBCOD 가수분해에 따른 난분해성물질 전환율 (-), b_H = 호기성 미생물 자가분해 상수 (/d), O = 용존산소 농도

Table 1. Variables and parameters of ordinary heterotrophic microbial growth-decay kinetic model categorized for purposes in numerical simulation and fitting analysis (Lee et al., 2014)

Category	Symbol	Definition	Text in the code*	Remarks
Dependent variables	X_H	Microorganism concentration [mgCOD/L]	y(1)	
	S_S	Readily biodegradable COD [mgCOD/L]	y(2)	
	X_S	Slowly biodegradable COD [mgCOD/L]	y(3)	
State variables	OUR	Oxygen Uptake Rate [mgO/L/d]	OURate	Data for fitting analysis
Initial conditions	$X_{H,0}$	X_H at t=0 [mgCOD/L]	y0(1)	
	$S_{S,0}$	S_S at t=0 [mgCOD/L]	y0(2)	Fitting parameters
	$X_{S,0}$	X_S at t=0 [mgCOD/L]	y0(3)	
Kinetic parameters	μ_H	Max. specific growth rate [/d]	mu_h	
	K_S	Half saturation coeff. for growth [mgCOD/L]	k_s	
	k_H	Max. specific hydrolysis rate [/d]	k_h	Fitting parameters
	K_X	Half saturation coeff. for hydrolysis [-]	k_x	
Stoichiometric constants	b_H	Heterotrophic decay rate [/d]	b_h	Fixed parameters
	Y_H	Heterotrophic yield coefficient [-]	y_h	Fixed parameters
	f_p	Biomass fraction yielding fragments [-]	f_p	

* See the numerical code in Appendices 1~3

(mgO/L), 그리고 $OUR =$ 산소소모율 (mgO/L/d). 본 연구의 베이지안 보정 기법 적용 시, 배치 실험에서 획득된 종속영양미생물의 산소호흡율(OUR_H) 데이터를 실험-모델 접합을 위한 상태 변수로 활용하였으며, 미생물 성장 동역학 모델 매개변수 4개(μ_H, K_S, k_H, K_X) 그리고 각 종속변수 초기값 3개($X_{H,0}, S_{S,0}, X_{S,0}$)를 데이터-모델 접합(Fitting) 계수로 활용하였다. 미생물 사멸 동역학 계수 b_H 및 화학양론 계수 Y_H, f_p 는 불확실성이 크지 않으므로 ASM1 모델에서 제시된 대푯값(0.62, 0.67, 0.08)을 사용하였다(Henze et al., 1987; Lee et al., 2014) (Table 1; Appendix 2).

2.4 응집 동역학 모델

본 연구의 응집 공정(반응) 동역학에 대한 베이지안 보정 기법 적용성 평가를 위하여, 이균집(Two-Class) 응집 동역학 모델을 적용하였다(Lee et al., 2011). 물환경에서 생체고분자 물질(EPS) 매개되는 응집은 응집핵(Flocculi)과 응집체(Floc) 이균집 입도분포를 형성하는 것으로 알려져 있다(Kim and Lee, 2020; Lee et al., 2012). 따라서, 본 연구는 입경이 고정된($\sim 5 \mu m$)을 가진 응집핵(Flocculi)과 변화(성장)하는 입경을 가진($10 \sim 500 \mu m$) 응집체(Floc)로 구성된 이균집입자계체군수지식(Two Class Population Balance Equation; TCPBE)을 적용하였다. TCPBE는 응집핵 농도(dN_P/dt), 응집체 농도(dN_F/dt) 및 플록에 결합된 응집핵의 총 농도(dN_T/dt)의 시간 변화율을 예측하는 연립 미분방정식 형태로 구성된다(Equation 5~7). TCPBE는 응집체에 결합된 응집핵의 수(NC)를 응집체 크기 지표로 사용하였으며, 충돌 빈도 인자(β)와 응집체 파괴 동역학 상수(a_F)에 대한 커널은 수체 전단력(G) 및 응집체 크기와 상관한 공식을 사용하였다(Equation 8~9)(Burd and Jackson, 2002; Lee et al., 2011). TCPBE 연립미분방정식은 Matlab (The MathWorks Inc., USA) 내장함수를 활용하여 적분되었으며, 베이지안 보정 기법과 연계하여 다수 모델 매개변수 조합에 대하여 연립미분방정식을 반복적으로 연산하였다. TCPBE 연립미분방정식 계산을 위한 Matlab 코드는 Appendix 4~6에 수록하였다.

$$\frac{dN_P}{dt} = -\frac{1}{2} a_{PP} \beta_{PP} N_P N_P \left(\frac{N_C}{N_C - 1} \right) - a_{PF} \beta_{PF} N_P N_F + f N_C a_F N_F \tag{5}$$

$$\frac{dN_F}{dt} = \frac{1}{2} a_{PP} \beta_{PP} N_P N_P \left(\frac{1}{N_C - 1} \right) - \frac{1}{2} a_{FF} \beta_{FF} N_F N_F + a_F N_F \tag{6}$$

$$\frac{dN_T}{dt} = \frac{1}{2} a_{PP} \beta_{PP} N_P N_P \left(\frac{N_C}{N_C - 1} \right) + a_{PF} \beta_{PF} N_P N_F - f N_C a_F N_F \tag{7}$$

$$\beta_{ij} = \frac{1}{6} (D_i + D_j)^3 G \tag{8}$$

$$a_F = E_b G \left(\frac{D_F - D_p}{D_p} \right)^b \left(\frac{\mu G}{F_y / D_F^2} \right)^q \tag{9}$$

여기서, D_P 와 D_F 는 각각 응집핵, 응집체의 평균 직경을 나타낸다. f 는 응집체 파괴에 의해 생성된 응집핵의 분율이다. a_{PP}, a_{PF}, a_{FF} 는 각각 응집핵-응집핵, 응집핵-응집체, 응집체-응집체 충돌 시 최종 부착되는 충돌 효율을 나타내며, 일반적으로 모델 보정 매개변수로 사용된다. 충돌 빈도 인자(β)는 수체 난류에 따른 입자 충돌의 확률을 나타내며, 아래첨자 i 와 j 는 응집체(P) 또는 응집체(F)를 나타낸다. E_b 는 응집체 파괴 동역학의 효율, μ 는 유체의 절대 점도, G 는 전단 속도, F_y 는 응집체 전단 강도, p 와 q 는 실험적으로 결정된 경험적 매개변수이다. 본 연구의 베이지안 보정 기법 적용 시, 응집 동역학 실험에서 획득된 응집체 평균입경(D_F) 및 부피분율(V_F)을 실험-모델 접합을 위한 상태 변수로 활용하였으며, 충돌 효율 인자(a_{PP}, a_{PF}, a_{FF}), 응집체 프랙탈 상수 (n_f), 응집체 파괴 동역학의 경험적 매개변수(E_b, p, q), 응집체 파괴에 의해 생성된 응집핵 분율(f)을 데이터-모델 접합 매개변수로 활용하였다. 이와 더불어, 응집핵의 초기 부피분율($V_{P,0}$)을 실험-모델 접합 매개변수로 활용하여 TCPBE 연립미분방정식의 초기조건($N_{P,0}, N_{F,0}, N_{T,0}$)을 산출하였다(Table 2; Appendix 5). 불확실성이 비교적 작은 응집 동역학 매개변수(F_y)는 기존 문헌의 대푯값(1.0×10^{-10})을 적용하였으며 (Winterwerp and van Kesteren, 2004; Lee et al., 2011), 반응 조 유체 흐름 및 입자 특성 매개변수들은 표준 및 실측값($\mu = 1.0 \times 10^{-6}, G = 95, D_p = 5.0 \times 10^{-6}, \rho_p = 2500, D_{F,0} = 15.0 \times 10^{-6}, C_{mass} = 0.12$)을 활용하였다(Table 2).

2.5 베이지안 보정 기법

베이지안 보정 기법은 베이지안 추론(Bayesian Inference)을 기반하는 확률론적 데이터-모델 보정 방법으로써, 관측 데이터-모델 결과 간 최적 보정(Fitting) 매개변수를 확률론적으로 추론하는 방법이다. 베이지안 보정 기법은 알려진 데이터를 기반으로 미지의 사후 분포를 추론하는 방법론이다(Equation 10). 베이지안 보정 기법은 비선형 또는 선형 관계를 설명할 수 있으며, 모델의 불확실성 정량 평가할 수 있는 장점을 가진다(Beckers et al. 2020).

$$P(X | \tilde{y}) = P(X) \cdot L(X | \tilde{y}) \tag{10}$$

본 연구에서는 베이지안 추론을 기반으로 하는 DREAM (Differential Evolution Adaptive Metropolis) 소프트웨어 패키지(<https://faculty.sites.uci.edu/jasper/software/>)를 사용하여 미생물 성장-사멸 동역학 및 응집 동역학 모델 매개변수를 추정하였다. 베이지안 보정의 결과는 실험 및 관측 데이터에 대하여 모델 접합(Fitting) 매개변수의 불확실성을 반영하여 매개변수의 사후 확률 분포를 나타낼 수 있다. 매개변수의 사후 분포는 사전 분포와 실험/관측 데이터(e.g., 미생물 성장-사멸 동역학 실험의 산소호흡률, 응집 동역학 실험의 응집체 크기 및 분율)를 고려한 매개변수의 확률(빈도) 함수로부터 계산된다. 본 연구는 베이지안 보정 기법의 적용에 초점을 두었으며, 자세한 수학적 알고리즘은 DREAM 개발자가 출간한 연구논문에 수록되어 있다(Vrugt, 2016).

Table 2. Variables and parameters of the two-class flocculation kinetic model categorized for purposes in numerical simulation and fitting analysis (Lee et al., 2011)

Category	Symbol	Definition	Text in the code*	Remarks
Dependent variables	N_p	Number concentration of flocculi [m^3]	y(1)	
	N_f	Number concentration of flocs [m^3]	y(2)	
	N_T	Number conc. of flocculi bound in flocs [m^3]	y(3)	
State variables	D_f	Mean diameter of flocs [m]	fdia	Data for fitting analysis
	V_f	Volume fraction of flocs [-]	ffrac	
Initial conditions	$N_{p,0}$	N_p at t=0	y0(1)	Dependent on $V_{p,0}$
	$N_{f,0}$	N_f at t=0	y0(2)	
	$N_{T,0}$	N_T at t=0	y0(3)	
	$V_{p,0}$	Volume fraction of primary particles at t=0 [-]	pfrac_ini	Fitting parameters
Kinetic parameters	α_{pp}	collision efficiency factor of flocculi [-]	alpha1	Fitting parameters
	α_{pf}	collision efficiency factor b/t flocculi & flocs [-]	alpha2	
	α_{ff}	collision efficiency factor of flocs [-]	alpha3	
	n_f	Fractal dimension of floc [-]	nfrac	
	E_b	Breakage coefficient [$s^{0.5}/m$]	brkEb	
	p	Empirical parameter of breakage kinetics [-]	brk_p	
	q	Empirical parameter of breakage kinetics [-]	brk_q	
	f	Fraction of flocculi by breakage [-]	brk_f	
Flow conditions	F_y	Yield strength of flocs [$10^{-10}Pa$]	brk_fy	Fixed parameters
	μ	Absolute viscosity of the fluid [N s/m ²]	mu	Fixed parameters
	G	Fluid turbulent shear rate [1/s]	shear	
Particle properties	D_p	Mean diameter of primary particles [m]	dp	Fixed parameters
	ρ_p	Density of primary particles [g/L]	pdens	
	$D_{f,0}$	Mean diameter of flocs at t=0 [m]	df_ini	
	C_{mass}	Mass concentration of particles [g/L]	pconc	

* See the numerical code in Appendices 4~6

3. Results and Discussion

3.1 최적 접합 매개변수 추적 알고리즘

본 연구의 베이지안 보정 기법은 DREAM 소프트웨어 (Vrugt, 2016)를 사용하였으며, 특별히 DREAM_(ZS) 알고리즘을 사용하여 최적 접합 매개변수 추적하였다. Fig. 3(a)와 같이, DREAM_(ZS) 알고리즘은 병렬연산(Parallel Computation)을 활용하여 1회 계산 배치에 독립적인 3개 매개변수 조합에 대하여 시뮬레이션을 동시에 수행하여 연산 시간을 단축하였다. 각 회차 계산 배치의 연산 결과를 바탕으로 최적 접합에 더 근접한 매개변수를 도출하여 후속 연산에 활용하므로, 시뮬레이션 결과가 점진적으로 최적 접합에 도달하게 된다. 본 연구의 미생물 성장-사멸 동역학 $X_{H,0}$ 매개변수 경우, 대략 1,000회 계산 배치 수행 후 최적값인 49.2 mg/L에 도달하는 것으로 나타났다(Fig. 3a). 또한, 계산 배치가 진행되면서 더 이상 발산하지 않고 안정화되는 여부(Convergence)를 판단하기 위하여 R-diagnostic 인자를 바탕으로 결정된다 (Gelman and Rubin, 1992). 본 연구 미생물 성장-사멸 동역학 테스트의 경우, 12,000회 계산 배치 후 안정화에 도달하는 것으로 나타났다(Fig. 3b).

3.2 미생물 성장-사멸 동역학

미생물 성장-사멸 동역학 배치 실험을 수행하여 획득한 산소소흡율(OUR) 곡선은 Fig. 4 같이 나타나며, 실험초기 호기성 미생물의 RBCOD 분해에 의해 OUR의 급격한 증가가 나타나며 RBCOD 분해가 종료되는 시점에 OUR이 급격히 하락하며, 이후 SBCOD 및 미생물 자가분해(사멸)에 따라 OUR이 일정하게 유지되는 경향을 가진다. 미생물 성장-사멸 동역학 실험의 이와 같은 OUR 변화 경향을 활용하여 하수에 포함된 RBCOD 농도($S_{S,0}$)를 계산할 수 있으며, 또한 활성 슬러지의 호기성 미생물 농도($X_{H,0}$) 계산에 활용할 수 있다 (Lee et al., 2014; Wentzel et al., 1995).

기존의 RBCOD 및 호기성 미생물 농도 계산은 연립미분방정식을 간략화, 실험 데이터에 선형화된 모델을 최적 접합(Fitting)하는 해석적 계산법(Analytical Methods)을 활용하였으나, 본 연구는 베이지안 보정 기법을 활용하여, 비선형 연립미분방정식 변경 없이, 사용자가 정의한 모델 매개변수의 사전 분포(Prior Distribution)에 대한 결과, 즉 사후 분포(Posterior Distribution)를 확률론적인 기법으로 산출하게 된다. Table 3과 같이, 사용자가 정의한 일정 범위를 가진 모델 매개변수의 사전 분포(Prior Distribution) 값(수천~수만 개)

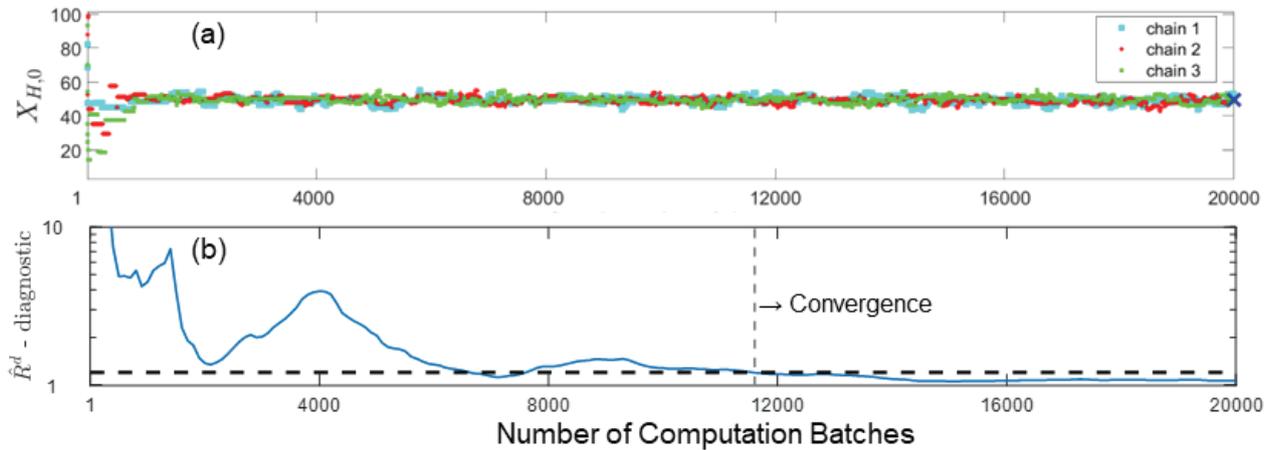


Fig. 3. (a) Trace plot of sampled values of a fitting parameter ($X_{H,0}$ of the microbial growth-decay kinetics), and (b) Evolution of the R convergence diagnostic of Gelman and Rubin (1992).

Table 3. Prior distribution and posterior mean of model-data fitting parameters of microbial growth-decay kinetics used in Bayesian calibration

Parameter	Definition	Prior distribution	Posterior mean
μ_H	Heterotrophic max. specific growth rate [1/d]	[0.1~10.0]	4.98
K_S	Half saturation coeff. for heterotrophs [mgCOD/L]	[0.001~0.10]	0.047
k_h	Max. specific hydrolysis rate [1/d]	[0.1~10.0]	5.61
K_X	Half saturation coeff. for hydrolysis [-]	[0.05~5.0]	3.09
$X_{H,0}$	Concentration of heterotrophs [mgCOD/L]	[1~100]	49.2
$S_{S,0}$	Readily biodegradable COD [mgCOD/L]	[10~100]	70.6
$X_{S,0}$	Slowly biodegradable COD [mgCOD/L]	[50~500]	322.8

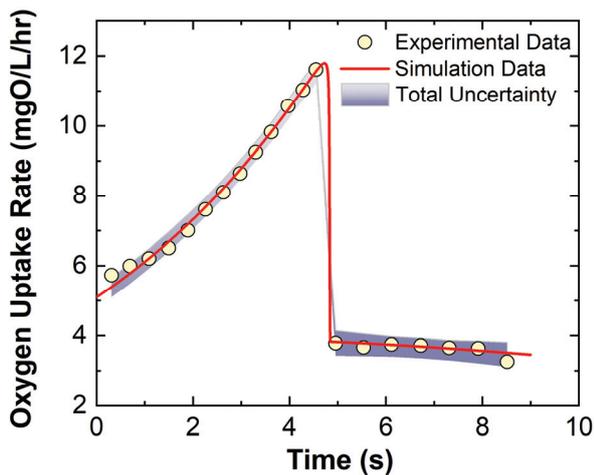


Fig. 4. Time series of oxygen uptake rate measured (dotted), best-fit simulation (red line), and 95% simulation uncertainty intervals due to total uncertainty (gray band).

들에 대하여 연립미분방정식을 연산하고 데이터-모델 접합도를 평가함으로써 사후 분포(Prior Distribution) 값들을 얻

게 된다. 본 연구에서는 번거로운 모델의 선형화, 단순화 과정 없이, 호기성 미생물 성장 동역학 모델의 주요 매개변수인 하수 RBCOD 농도($S_{S,0}$), 활성슬러지 미생물 농도($X_{H,0}$)를 평가할 수 있었다. 다만 베이지안 보정 기법은 수천~수만 개 연립미분방정식 풀이 연산을 해야 하지만, 현재의 가정용 PC 수준으로도 10분 내로 결과를 도출할 수 있다.

모델 매개변수의 사후 분포는 Fig. 5와 같이 일반적으로 정규분포의 형태를 가지지만, 일부 민감도가 낮은 매개변수는 명확한 정규분포를 띄지 않고 평탄한 형태를 가진다. 이와 같이 비교적 평탄한 사후 분포를 가지는 매개변수들은 모델 연산 결과에 큰 영향을 미치지 아니하므로 데이터-모델 접합 작업에서 제외할 수 있다. 또한, 모델 매개변수들 중 상관성이 매우 높은 경우가 발생하며, 예를 들어, $\mu_H \sim X_{H,0}$ 또는 $k_h \sim S_{S,0}$ 의 경우, 사용자의 판단에 따라 상응하는 두 개 매개변수 중 하나를 취사선택하여 데이터-모델 보정 작업의 복잡성을 줄일 수 있다. 이와 같이 베이지안 보정 기법은 최적 매개변수의 도출 기능뿐만 아니라, 모델 예측 결과에 대한 매개변수의 영향과 매개변수들 간 상관성에 대한 정보를 추가 제공할 수 있다. 최근 대체인 머신러닝, 인공지능 기법들과 비교하여서도, 베이지안 보정 기법은 매개변수 특성 평가에 관한 특별한 장점을 가지고 있다.

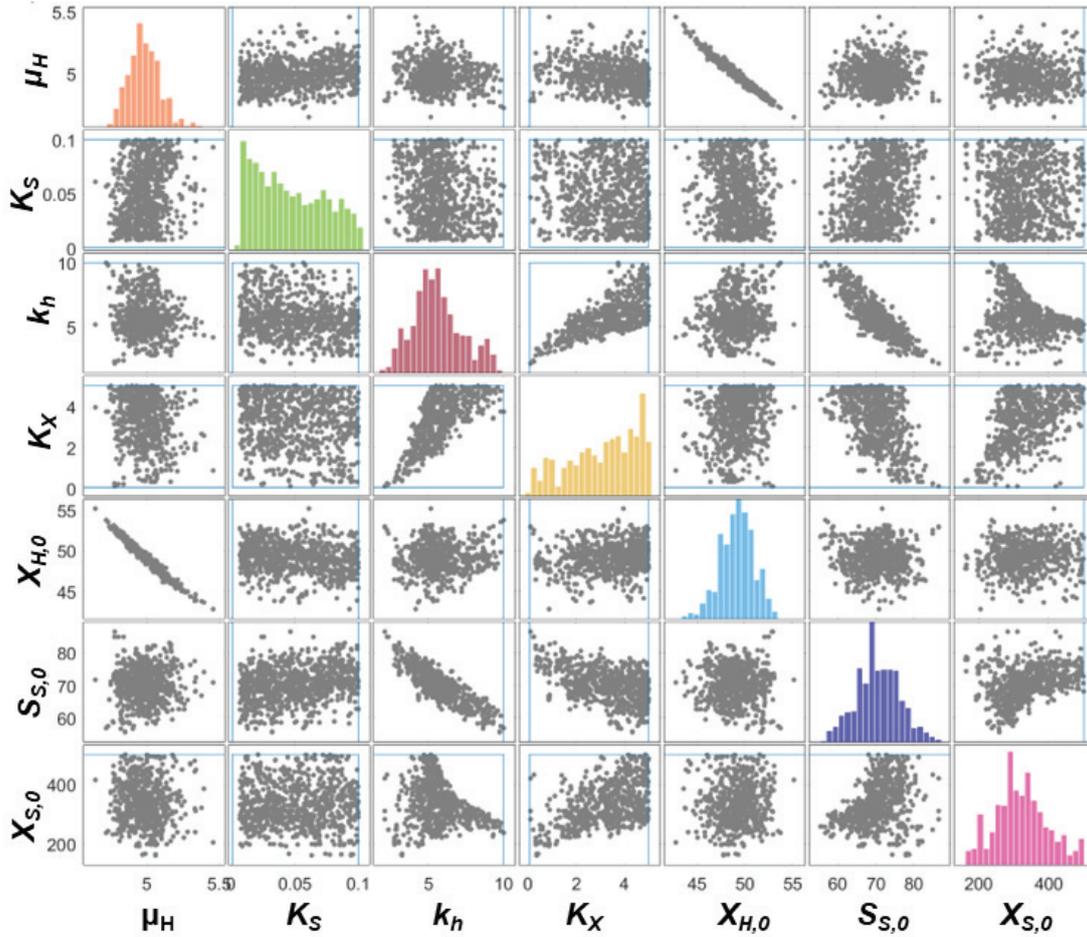


Fig. 5. Histogram of posterior distribution of microbial growth-decay kinetic parameters used in Bayesian calibration and correlation plots between different parameters

3.3 응집 동역학

본 연구에서는 응집핵-응집체 이균집 응집 동역학 모델(TCPBE)을 적용하였으므로, 응집체 크기(D_f)와 전체 응집핵-응집체 농도 대비 응집체 분율(V_f)을 실험 데이터-모델 보정 작업에 활용하였다. 응집 동역학 실험 동안, 응집을 통하여 응집핵의 상대적 분율을 감소시키고, 응집체의 크기와 분율은

지속하여 커지며 더 이상의 변화가 없는 평형 상태에 도달한다(Fig. 6). TCPBE는 이와 같은 응집핵-응집체의 변화 경향을 모사할 수 있으며, 베이저안 보정 기법을 적용하여 실험 데이터-모델 예측치 간 최적 접합의 결과를 도출할 수 있었다(Table 4).

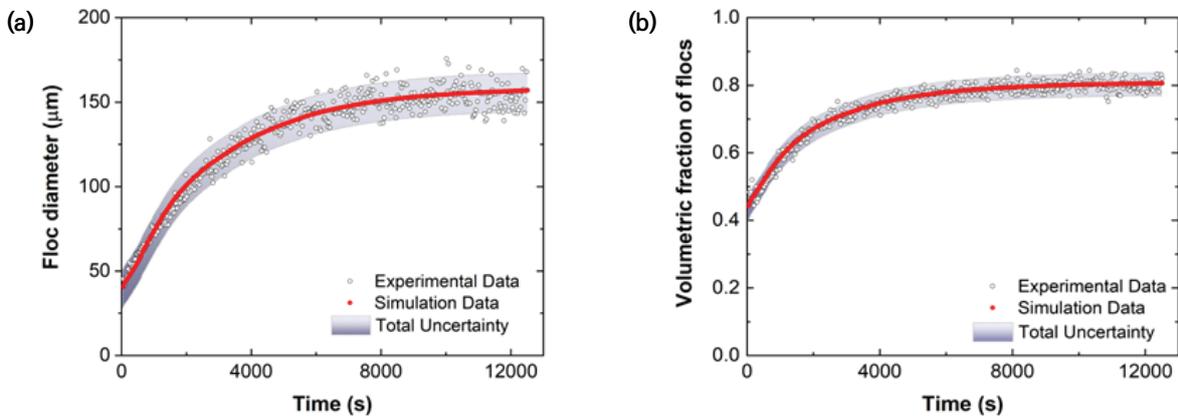


Fig. 6. Time series of (a) floc diameter and (b) volumetric fraction measured (dotted), best-fit simulation (red line), and 95% simulation uncertainty intervals due to total uncertainty (gray band).

Table 4. Prior distribution and posterior mean of model-data fitting parameters of two-class flocculation kinetics used in Bayesian calibration

Parameter	Definition	Prior distribution	Posterior mean
α_{PP}	Collision efficiency factor of flocculi [-]	[0.0~0.25]	0.008
α_{PF} α_{FF}	Collision efficiency factor of floc [-]	[0.0~1.00]	0.997
n_f	Fractal dimension of floc [-]	[2.0~2.6]	2.37
E_b	Breakage coefficient [$s^{0.5}/m$]	$[1.0 \times 10^{-5} \sim 7.5 \times 10^{-4}]$	4.0×10^{-5}
p	Empirical parameter of breakage kinetics [-]	[0.1~2.5]	0.828
q	Empirical parameter of breakage kinetics [-]	[0.15~1.5]	0.265
f	Fraction of flocculi by breakage [-]	[0.0~1.0]	0.009
$V_{P,0}$	Initial volumetric fraction of primary particle [-]	[0.5~1.0]	0.920

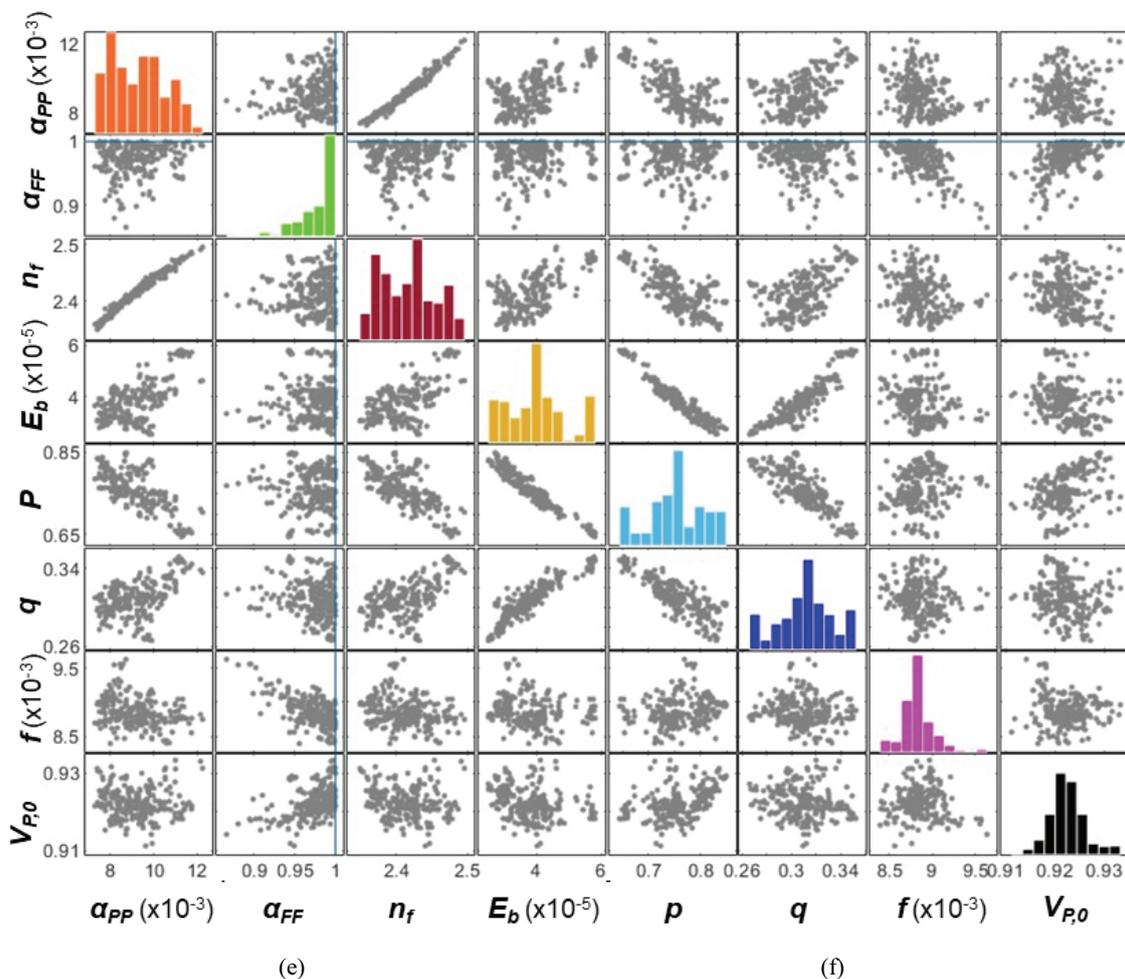


Fig. 7. Histogram of posterior distribution of flocculation kinetic parameters used in Bayesian calibration and correlation plots between different parameters.

앞서 소개한 미생물 성장-사멸 동역학의 경우와 유사하게, 사용자가 정의한 응집 동역학 모델 매개변수의 사전 분포 (Prior Distribution)에 대한 결과(Posterior Distribution)를 확률론적인 기법으로 산출하게 된다. 데이터-모델 접합을 위한 다수 매개변수들에 대한 일괄적인 평가는 어려운 작업이며, 인위적인 시행착오(Trials and Errors)로 산출하기는 거의 불가능하다. 베이지안 보정 기법은 수천, 수만회의 반복 작업을 컴퓨

터를 활용하여 수행하며, 그 결과를 통계적으로 해석함으로써 비교적 손쉽게 데이터-모델 보정 작업을 수행하게 된다.

응집 동역학 모델 매개변수의 사후 분포(Posterior Distribution)는 대체로 정규분포를 따르며, 특별히 응집체 충돌 효율 계수(α_{FF})는 사용자가 정의한 사전 분포의 상단 경계에 최빈값이 형성되는 것으로 나타났다(Fig. 7). α_{FF} 는 물리적으로 1을 넘을 수 없으므로, 매개변수의 사전분포 범위 설정에서 현실

의 물리적인 범위 내로 한정하였기 때문에, 사후분포는 상단 경계 1을 넘지 않고 그 주변으로 집중되어 있었다. α_{FF} 와 응집체 프랙탈 차원(형상) 계수(n_f)는 강한 상관성을 나타내며, 이들 매개변수가 공통으로 응집 동역학을 증대시키는 것으로 파악할 수 있었다. 또한, 응집체 파괴와 관련된 세 개의 매개변수들(E_b , p , q) 간 뚜렷한 음 또는 양의 상관성을 나타내고 있다. 이와 같이, 모델 매개변수들의 상관성을 파악함으로써, 모델의 구조적 특성 이해를 도모하고 사용자의 모델 적용의 방향에 따라 매개변수를 취사 선택할 여지를 제공할 수 있다.

4. Conclusion

최근 각광을 받고 있는 인공지능 기반 블랙박스 모델들에 비하여, 연립미분방정식 기반 수학적 예측 모델은 생물-물리-화학적 공정(반응) 동역학을 반영하며 사전 데이터가 존재하지 않더라도 모델을 구동할 수 있는 장점을 가지고 있다. 하지만, 다양한 생물-물리-화학적 공정(반응) 동역학이 얽혀 있으며 수많은 매개변수를 포함하는 연립미분방정식을 풀어 내기는 쉽지 않다. 때로는 체계적인 매개변수 평가 없이, 인위적인 시행착오를 통하여 얻어진 모델 매개변수 변수들을 활용하기도 한다.

본 연구에서는 베이지안 보정 기법을 미생물 성장-사멸 동역학과 응집 동역학 실험 및 모델에 대하여 적용하였으며, 베이지안 보정 기법과 연계한 수학적 예측 모델의 새로운 가능성을 제시하였다. 본 연구논문에서 제시된 베이지안 보정 기법을 활용하여, 연구자가 수립한 다양한 생물-물리-화학적 공정(반응) 동역학 모델의 매개변수 보정 작업을 수월하게 진행할 수 있으리라 판단된다.

현재의 컴퓨팅 기술의 발전을 고려하면, 베이지안 보정 기법은 복잡한 생물-물리-화학적 공정(반응) 동역학 모델에 대한 적용뿐만 아니라, 항후 반응-이동 현상을 고려한 3차원 시공간 예측모델에도 적용가능하리라 판단된다. 베이지안 보정 기법은, 컴퓨팅 기술의 발전과 함께, 물환경 수질예측에 주로 활용되는 대규모, 다차원 수리-수질 예측 모델(EFDC, Delft3D, MIKE 등)에도 적용할 수 있으리라 기대된다. 또한, 최근 4차 산업혁명의 기조에 따라 국내외 상하수처리 플랜트의 디지털 트윈 구현이 활발히 진행되고 있다. 본 연구에서 제시한 베이지안 보정 기법과 상하수처리 물리-생물-화학적 공정 동역학 모델을 연계한다면, 인공지능-머신러닝 등 데이터 집약형 예측 모델에 대비하여, 저비용 디지털 트윈 구현이 가능할 것으로 기대된다.

Acknowledgement

이 논문은 2021년도 한국연구재단 이공분야 기초연구사업(NRF-2020R1I1A3A04036895)의 지원을 받아 수행됨.

References

- Beckers, F., Heredia, A., Noack, M., Nowak, W., Wieprecht, S., and Oladyskhin, S. (2020). Bayesian calibration and validation of a large-scale and time-demanding sediment transport model, *Water Resources Research*, 56(7), e2019WR026966. <https://doi.org/10.1029/2019WR026966>
- Burd, A. B. and Jackson G. A. (2002). Modeling steady-state particle size spectra, *Environmental Science and Technology*, 36(3), 323-327. <https://doi.org/10.1021/es010982n>
- Choi, Y., Baek, S., Kim, J., Choi, J., Hur, J., Lee, T., Park, C., and Lee B. J. (2017). Characteristics and biodegradability of wastewater organic matter in municipal wastewater treatment plants collecting domestic wastewater and industrial discharge, *Water*, 9(6), 409. <https://doi.org/10.3390/w9060409>
- Duarte, M. S., Martins, G., Oliveira, P., Fernandes, B., Ferreira, E. C., Alves, M., Lopes, F., Pereira, M. A., and Novais, P. (2024). A review of computational modeling in wastewater treatment processes, *ACS ES&T Water*, 4(3), 784-804. <https://doi.org/10.1021/acsestwater.3c00117>
- Gelman, A. G. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences, *Statistical Science*, 7(4), 457-472. <https://doi.org/10.1214/ss/1177011136>
- Gujer, W., Henze, M., Mino, T., and van Loosdrecht, M. (1999). Activated sludge model No. 3, *Water Science and Technology*, 39(1), 183-193. [https://doi.org/10.1016/S0273-1223\(98\)00785-9](https://doi.org/10.1016/S0273-1223(98)00785-9)
- Henze, M., Grady, C. P. L., Gujer, W., Marais, G. v. R., and Matsuo, T. (1987). *Activated sludge model No. 1*, IAWPRC, London, England.
- Jeppsson, U. (1996). *Modelling aspects of wastewater treatment processes*, PhD Dissertation, Lund Institute of Technology, Lund, Sweden.
- Kim, J. I. and Lee, B. J. (2020). Investigation on flocculi-floc interaction and flocculation in extracellular polymeric substances, ionic species and clay-containing suspension, *Journal of Korean Society on Water Environment*, 36(3), 185-193. <https://doi.org/10.15681/KSWE.2020.36.3.185>
- Lee, B. J., Fettweis, M., and Toorman, E., and Molz, F. J. (2012). Multimodality of a particle size distribution of cohesive suspended particulate matters in a coastal zone, *Journal of Geophysical Research: Oceans*, 117, C03014. <https://doi.org/10.1029/2011JC007552>
- Lee, B. J., Toorman, E., Molz, F. J., and Wang, J. (2011). A two-class population balance equation yielding bimodal flocculation of marine or estuarine sediments, *Water Research*, 45(5), 2131-2145. <https://doi.org/10.1016/j.watres.2010.12.028>
- Lee, B. J., Wentzel, M. C., and Ekama G. A. (2006). Measurement and modelling of ordinary heterotrophic organism active biomass concentrations in anoxic/aerobic activated sludge mixed liquor, *Water Science and Technology*, 54(1), 1-10. <https://doi.org/10.2166/wst.2006.365>

- Lee, B. J., Wentzel, M., Ekama, G. A., Choi, Y., and Choi, J. (2014). Measurement of ordinary heterotrophic organism active biomass in activated sludge mixed liquor: Evaluation and comparison of the quantifying techniques, *Environmental Engineering Research*, 19(1), 91-99. DOI: <https://doi.org/10.4491/eer.2014.19.1.091>
- Van Loosdrecht, M. C. M., Lopez-Vazquez, C. M., Meijer, S. C. F., Hooijmans, C. M., and Brdjanovic, D. (2015). Twenty-five years of ASM1: Past, present and future of wastewater treatment modelling, *Journal of Hydroinformatics*, 17(5), 697-718. <https://doi.org/10.2166/hydro.2015.006>
- van Mourik, S., ter Braak, G., Stigter, H., and Molenaar, J. (2014). Prediction uncertainty assessment of a systems biology model requires a sample of the full probability distribution of its parameters, *PeerJ*, 2, e433. <https://doi.org/10.7717/peerj.433>
- van Turnhout, A. G., Kleerebezem, R., and Heimovaara, T. J. (2016). A toolbox to find the best mechanistic model to predict the behavior of environmental systems, *Environmental Modelling & Software*, 83, 344-355. <https://doi.org/10.1016/j.envsoft.2016.05.002>
- Villacorte, L. O., Ekowati, Y., Neu, T. R., Kleijn, J. M., Winters, H., Amy, G., Schippers, J. C., and Kennedy, M. D. (2015). Characterisation of algal organic matter produced by bloom-forming marine and freshwater algae, *Water Research*, 73, 216-230. <https://doi.org/10.1016/j.watres.2015.01.028>
- Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation, *Environmental Modelling & Software*, 75, 273-316. <https://doi.org/10.1016/j.envsoft.2015.08.013>
- Wentzel, M., Mbewe, A., and Ekama, G. (1995). Batch test for measurement of readily biodegradable COD and active organism concentrations in municipal waste waters, *Water SA*, 21(2), 117-124.
- Winterwerp, J. C. and van Kesteren, W. G. M. (2004). *Introduction to the physics of cohesive sediment in the marine environment*, Elsevier, Amsterdam, the Netherlands.

Appendix 1. Script for connecting the microbial growth-decay kinetic model and DREAM software

```

% -----
% Bayesian Calibration of the microbial growth-decay kinetics, with the DREAM software
%
% 2024.01.15 at SangJu, Kyungpook National University, written by BJ Lee
% Connecting the microbial growth-decay kinetic model and DREAM(c) software (Jasper A. Vrugt)
% Reference:
% Vrugt, J.A. (2016) Markov chain Monte Carlo simulation using the DREAM software package:
% Theory, concepts, and MATLAB implementation,
% Environmental Modelling & Software, 75, 273-316.
% -----

%% Problem settings defined by user
DREAMPar.d = 7; % Dimension of the problem
DREAMPar.lik = 11; % Likelihood 23 (limits of acceptability)

%% Initial sampling and parameter ranges
Par_info.initial = 'latin'; % Latin hypercube sampling
Par_info.boundhandling = 'reflect'; % Explicit boundary handling: reflection
% mu_h k_s k_h k_x XH_0 Ss_0 Xs_0
Par_info.min = [0.10 0.001 0.10 0.050 1.000 10.0 50.00]; % Lower bound parameters
Par_info.max = [10.0 0.100 10.0 5.000 100.0 100.0 500.0]; % Upper bound parameters

%% Define name of function (.m file) for posterior exploration
Func_name = 'OHOkineitics_0D_Increment';

%% Load calibration data vector (summary metrics with which simulations are compared)
MeasData = csvread('batchtest_results.csv');
Meas_info.Y = MeasData(1:end,2);

%% Optional settings (limits of acceptability)
options.parallel = 'yes'; % Run each chain on a different core
options.save = 'yes';
options.restart = 'no';
%options.modout = 'yes';

%% Define method to use {'dream','dream_zs','dream_d','dream_dzs'}
method = 'dream_zs';
switch method
case {'dream','dream_d'}
    DREAMPar.N = 10; % Number of Markov chains
    DREAMPar.T = 200; % Number of generations
case {'dream_zs','dream_dzs'}
    DREAMPar.N = 3; % Number of Markov chains
    DREAMPar.T = 5000; % Number of generations
end
if strcmp(method,'dream_d') || strcmp(method,'dream_dzs')
    Par_info.steps = 200*ones(1,DREAMPar.d); % Number of discrete steps
end

%% Run DREAM package
[chain,output,FX,Z,logL] = DREAM_package(method,Func_name,DREAMPar,Par_info,Meas_info,options);

```

Appendix 2. Function for integrating the microbial growth-decay kinetics

```
function [ OHO_simdata ] = OHOkinetics_0D_Increment(x)
% -----
% Function of integrating the microbial growth-decay kinetics
%
% 2024.01.15 at SangJu, Kyungpook National University, written by BJ Lee
% Numerical Integration the microbial growth-decay kinetic model
% Reference:
% Jeppsson, U. (1996) Modelling aspects of wastewater treatment processes,
% PhD Dissertation, Lund Institute of Technology.
% -----

% ODE_TCPBE Master
% Ordinary Heterotrophic Organisms (OHO) Kinetic Parameters
prOHO.mu_h=x(1); prOHO.k_s=x(2); prOHO.k_h=x(3); prOHO.k_x=x(4);
prOHO.b_h=0.62; prOHO.y_h=0.667; prOHO.f_p=0.08;

% Initial Conditions
y0 = [x(5); x(6); x(7)];

% Time steps to print out
MeasData = csvread('batchtest_results.csv'); % OUR_H data from the batch respiration test
ts_obs = MeasData(1:end,1);

t = linspace(0,0.375,10000)';
t_hour = t.*24.0;

% ODE4 calculates the solution, based on a fixed time step.
% ODE4 can be found in the MATLAB "File Exchange".
y = ode4(@(t,y)ODE_OHO(t,y,prOHO),t,y0);

% ASM1 Model, assuming that SBCOD is hydrolyzed and topped up in RBCOD
OURate = ((1.0/prOHO.y_h-1.0)*prOHO.mu_h*y(:,2)./(prOHO.k_s+y(:,2)).*y(:,1))./24.0;

OHO_simdata=zeros(size(ts_obs,1),1);
for i=1:size(ts_obs,1)
    [d ix] = min(abs(t_hour-ts_obs(i)));
    OHO_simdata(i)=OURate(ix);
end
```

Appendix 3. Function of the ordinary differential equations of the microbial growth-decay kinetics

```
function dydt = ODE_OHO(t,y,prOHO)
% -----
% Function of the ordinary differential equations of the microbial growth-decay kinetics
%
% 2024.01.15 at SangJu, Kyungpook National University, written by BJ Lee
% Coupled ordinary differential equation of the microbial growth-decay kinetic model
% Reference:
% Jeppsson, U. (1996) Modelling aspects of wastewater treatment processes,
% PhD Dissertation, Lund Institute of Technology.
% -----

mu_h=prOHO.mu_h;      k_s=prOHO.k_s;      k_h=prOHO.k_h;      k_x=prOHO.k_x;
b_h=prOHO.b_h;      y_h=prOHO.y_h;      f_p=prOHO.f_p;

% ASM1 Model, assuming that SBCOD is hydrolyzed and topped up in RBCOD
dydt = [mu_h*y(2)/(k_s+y(2))*y(1)-b_h*y(1);...
        -(1.0/y_h)*mu_h*y(2)/(k_s+y(2))*y(1)+k_h*(y(3)/y(1))/(k_x+(y(3)/y(1)))*y(1);...
        +(1.0-f_p)*b_h*y(1)-(1.0/y_h)*k_h*(y(3)/y(1))/(k_x+(y(3)/y(1)))*y(1)];
end
```

Appendix 4. Script for connecting the two-class flocculation kinetic model and DREAM software

```

% -----
% Bayesian Calibration of the flocculation kinetic model, with the DREAM software
%
% 2024.01.15 at SangJu, Kyungpook National University, written by BJ Lee
% Connecting the flocculation kinetic model and DREAM(c) software (Jasper A. Vrugt)
% Reference:
% Lee, B.J. et al. (2011) A two-class population balance equation yielding bimodal flocculation of
% marine or estuarine sediments, Water Research, 45(5), 2131-2145.
% -----

%% Problem settings defined by user
DREAMPar.d = 8;           % Dimension of the problem
DREAMPar.lik = 11;        % Likelihood 23 (limits of acceptability)

%% Initial sampling and parameter ranges
Par_info.initial = 'latin'; % Latin hypercube sampling
Par_info.boundhandling = 'reflect'; % Explicit boundary handling: reflection
% alpha1 alpha2 nfrac brk_Eb brk_p brk_q brk_f dfloc_ini
Par_info.min = [5.0e-3 0.10 2.0 2.5e-5 0.10 0.10 0.00 0.500]; % Lower bound parameters
Par_info.max = [5.0e-2 1.00 2.6 7.5e-4 1.50 1.50 0.20 0.999]; % Upper bound parameters

%% Define name of function (.m file) for posterior exploration
Func_name = 'flockinetics_OD';

%% Load calibration data vector (summary metrics with which simulations are compared)
MeasData = csvread('01MNaCl_5mgLXG_proc_cut.csv'); % Temperature data are summary statistics

% Normalization: Min ~ Max
MeasData2=(MeasData(1:end,2)-min(MeasData(1:end,2)))/(max(MeasData(1:end,2))-min(MeasData(1:end,2)));
MeasData3=(MeasData(1:end,3)-min(MeasData(1:end,3)))/(max(MeasData(1:end,3))-min(MeasData(1:end,3)));
MeasData_spaceL = mean(MeasData2(end-4:end,1))*ones(20,1);
MeasData_space = 0.5*ones(20,1);
MeasData_spaceR = mean(MeasData3(1:5,1))*ones(20,1);

% Combine the Normalized TS Data
Meas_info.Y = vertcat(MeasData2,MeasData_spaceL,MeasData_space,MeasData_spaceR,MeasData3);

%% Optional settings (limits of acceptability)
options.parallel = 'yes'; % Run each chain on a different core
options.save = 'yes';
options.restart = 'no';

%% Define method to use {'dream','dream_zs','dream_d','dream_dzs'}
method = 'dream_zs';

switch method
case {'dream','dream_d'}
    DREAMPar.N = 10; % Number of Markov chains
    DREAMPar.T = 200; % Number of generations
case {'dream_zs','dream_dzs'}
    DREAMPar.N = 3; % Number of Markov chains
    DREAMPar.T = 200; % Number of generations
end

if strcmp(method,'dream_d') || strcmp(method,'dream_dzs')
    Par_info.steps = 200*ones(1,DREAMPar.d); % Number of discrete steps
end

%% Run DREAM package
[chain,output,FX,Z,logL] = DREAM_package(method,Func_name,DREAMPar,Par_info,Meas_info,options);

```

Appendix 5. Function for integrating the flocculation kinetics

```
function [ flocc_simdata ] = flocckinetics_OD(x)
% -----
%           Function of integrating the two-class flocculation kinetic model
%
% 2024.01.15 at SangJu, Kyungpook National University, written by BJ Lee
% Numerical Integration of the two-class flocculation kinetic model
% Reference:
% Lee, B.J. et al. (2011) A two-class population balance equation yielding bimodal flocculation of
% marine or estuarine sediments, Water Research, 45(5), 2131-2145.
% -----

% ODE_TCPBE Master
% Flocculation Kinetic Parameters
% Assuming that alpha2 (collision efficiency between floc & flocculi) is the same as alpha 1!!!
prFloc.alpha1=x(1); prFloc.alpha2=prFloc.alpha1; prFloc.alpha3=x(2); prFloc.nfrac=x(3);
prFloc.brk_eb=x(4); prFloc.brk_p=x(5); prFloc.brk_q=x(6); prFloc.brk_f=x(7);
pfrac_ini=x(8);

prFloc.df_ini=15.0e-6; prFloc.dp=5.0e-6; prFloc.shear=95.0; prFloc.brk_fy=1.0e-10;
prFloc.mu=1.0e-6; pconc=0.12; pdens=2500;

df_ini=prFloc.df_ini;
nc_ini=(df_ini/prFloc.dp)^prFloc.nfrac;
pntot=pconc/(pdens*1.0/6.0*pi^0*prFloc.dp^3);

% Initial Conditions
y0 = [pntot*pfrac_ini; pntot*(1.0-pfrac_ini)/nc_ini; pntot*(1.0-pfrac_ini)];

% Time steps to print out
MeasData = csvread('01MNaCl_5mgLXG_proc_cut.csv');
ts_obs = MeasData(1:end,1);
MeasData2=(MeasData(1:end,2)-min(MeasData(1:end,2)))/(max(MeasData(1:end,2))-min(MeasData(1:end,2)));
MeasData3=(MeasData(1:end,3)-min(MeasData(1:end,3)))/(max(MeasData(1:end,3))-min(MeasData(1:end,3)));

t0=0; tfinal=ts_obs(end,1); nsteps=tfinal+1;
t = linspace(t0,tfinal,nsteps);

% ODE4 calculates the solution, based on a fixed time step
% ODE4 can be found in the MATLAB "File Exchange".
y = ode4(@(t,y)ODE_TCPBE(t,y,prFloc),t,y0);

%ffrac should be volumetric fraction!!!
fdia=(y(:,3)/y(:,2)).^(1.0/prFloc.nfrac)*prFloc.dp*1.0e6;
fdia_m=fdia./1.0e6;
ffrac=(0.1667*pi^0*fdia_m.^3).*y(:,2)/((0.1667*pi^0*prFloc.dp^3)*y(:,1)+(0.1667*pi^0*fdia_m.^3).*y(:,2));

% Normalization: Min ~ Max
ts_simData2 = (fdia(ts_obs)-min(MeasData(1:end,2)))/(max(MeasData(1:end,2))-min(MeasData(1:end,2)));
ts_simData3 = (ffrac(ts_obs)-min(MeasData(1:end,3)))/(max(MeasData(1:end,3))-min(MeasData(1:end,3)));

ts_simData_spaceL = mean(MeasData2(end-4:end,1))*ones(20,1);
ts_simData_space = 0.5*ones(20,1);
ts_simData_spaceR = mean(MeasData3(1:5,1))*ones(20,1);

floc_simdata = vertcat(ts_simData2,ts_simData_spaceL,ts_simData_space,ts_simData_spaceR,ts_simData3);
```

Appendix 6. Function of the ordinary differential equations of the flocculation kinetics

```

function dydt = ODE_TCPBE(ty,prFloc)
% -----
% Function of integrating the two-class flocculation kinetic model
%
% 2024.01.15 at SangJu, Kyungpook National University, written by BJ Lee
% Coupled ordinary differential equation of the two-class flocculation kinetic model
% Reference:
% Lee, B.J. et al. (2011) A two-class population balance equation yielding bimodal flocculation of
% marine or estuarine sediments, Water Research, 45(5), 2131-2145.
% -----

alpha1=prFloc.alpha1;    alpha2=prFloc.alpha2;    alpha3=prFloc.alpha3;
dp=prFloc.dp;           nfrac=prFloc.nfrac;    shear=prFloc.shear;
brk_eb=prFloc.brk_eb;   brk_p=prFloc.brk_p;    brk_q=prFloc.brk_q;    brk_f=prFloc.brk_f;
brk_fy=prFloc.brk_fy;   mu=prFloc.mu;

%yp = [(1 - param.alpha*y(2))*y(1); (-1 + param.beta*y(1))*y(2)];
dydt = [(-2.0/3.0*alpha1*dp*dp*dp*shear*y(1)*y(1)*((y(3)/y(2))/((y(3)/y(2))-1.0))...
-1.0/6.0*alpha2*((y(3)/y(2))^(1.0/nfrac)+1.0)^3.0*dp*dp*dp*shear*y(1)*y(2)...
+brk_f*(y(3)/y(2))*brk_eb*shear*(y(3)/y(2))^(1.0/nfrac)-1.0)^brk_p...
*(mu*shear*(y(3)/y(2))^(2.0/nfrac)*dp*dp/brk_fy)^brk_q*y(2)];...
(+2.0/3.0*alpha1*dp*dp*dp*shear*y(1)*y(1)*(1.0/((y(3)/y(2))-1.0))...
-2.0/3.0*alpha3*(y(3)/y(2))^(3.0/nfrac)*dp*dp*dp*shear*y(2)*y(2)...
+brk_eb*shear*(y(3)/y(2))^(1.0/nfrac)-1.0)^brk_p...
*(mu*shear*(y(3)/y(2))^(2.0/nfrac)*dp*dp/brk_fy)^brk_q*y(2)];...
(+2.0/3.0*alpha1*dp*dp*dp*shear*y(1)*y(1)*((y(3)/y(2))/((y(3)/y(2))-1.0))...
+1.0/6.0*alpha2*((y(3)/y(2))^(1.0/nfrac)+1.0)^3*dp*dp*dp*shear*y(1)*y(2)...
-brk_f*(y(3)/y(2))*brk_eb*shear*(y(3)/y(2))^(1.0/nfrac)-1.0)^brk_p...
*(mu*shear*(y(3)/y(2))^(2.0/nfrac)*dp*dp/brk_fy)^brk_q*y(2)];

end

```