SUPPLEMENTARY DATA

Why are organic micropollutants not fully biotransformed? A mechanistic modelling approach to anaerobic systems

Lorena Gonzalez-Gil, Miguel Mauricio-Iglesias*, Marta Carballa, Juan M. Lema

Department of Chemical Engineering, School of Engineering, Universidade de Santiago de Compostela, Rúa Lope Gómez de Marzoa, E-15782 Santiago de Compostela, Spain

***** Corresponding author

E-mail addresses: *lorena.gonzalez@usc.es*, *miguel.mauricio@usc.es*, *juan.lema@usc.es*, *marta.carballa@usc.es*

Contents

S1. Batch anaerobic digesters performance

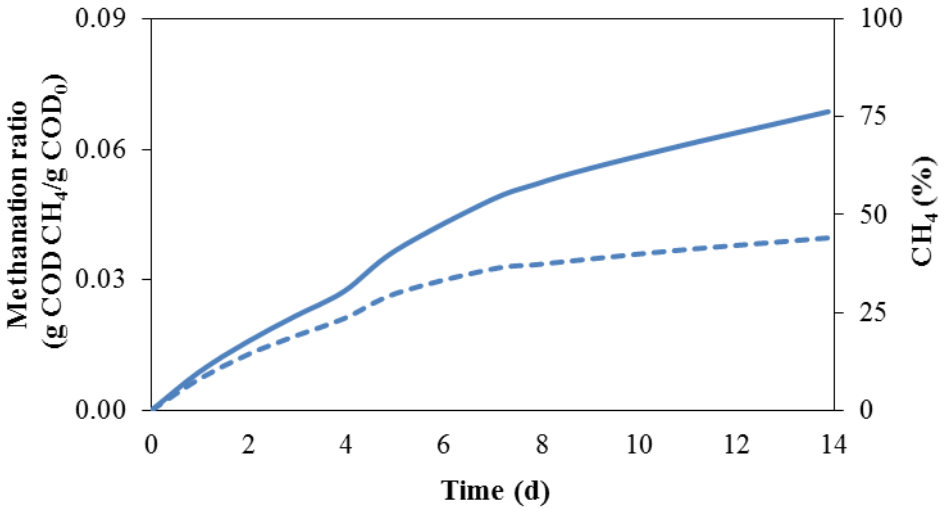

Figure S1. Performance of batch anaerobic digesters (Experiment 4) in terms of COD methanized (continuous line) and biogas composition (dashed line) inside the head-space of the bottles.

S2. Bootstrap method for model calibration

The bootstrap procedure was done as follows to determine the expected value and the confidence interval of the parameters:

i) Perform a reference parameter estimation using the trust-region-reflective algorithm as implemented in Matlab routine *lsqnonlin* to solve Eq. (S1).

$$\theta^* = \arg\min_{\theta} \sum_{k=1}^{m} \sum_{i=1}^{n} \left(C_{exp,k}(t_i) - C_{pred,k}(t_i, \theta) \right)^2 \quad \text{(S1)}$$

ii) Estimate the reference residuals (*e*), which are used to simulate the experimental error in the measurements. The residuals are assumed to have equal probability of realization (Eq. (S2)).

$$e = C_{exp,k}(t_i) - C_{pred,k}(t_i, \theta^*) \quad \text{(S2)}$$

iii) Generate synthetic data ($C^*_{exp,k}$) by randomly adding some of the reference residuals (random sampling with replacement) to the experimental data (Eq. (S3)).

$$C^*_{exp,k} = C_{exp,k} + e_j \quad \text{where } e_j \text{ are elements randomly sampled from } e \quad \text{(S3)}$$

iv) Estimate parameters solving Eq. (S4) for the set of generated synthetic data to find the set of parameters θ_j corresponding to iteration *j*.

$$\theta_j = \arg\min_{\theta} \sum_{k=1}^{m} \sum_{i=1}^{n} \left(C^*_{exp,k}(t_i) - C_{pred,k}(t_i, \theta) \right)^2 \quad \text{(S4)}$$

v) Iterate through steps iii) and iv). The mean and standard deviation of the distribution of θ_j were checked as indications of convergence (number of iterations, n_{It} = 200).

vi) It is assumed that for a large n_{It}, the expected value (Eq. (S5)) can be approximated as the mean of θ_j and the confidence interval for a significance α (in this work 0.05) of the calibrated parameters (Eq. (S6)) is estimated from the quantile function.

$$\hat{\theta} = \frac{1}{n_{It}} \sum_{j=1}^{nIt} \theta_j \quad \text{(S5)}$$

$$\hat{\theta}_{\alpha/2} = \left\{ \theta : \ Pr\left(\theta_j \le \hat{\theta}_{\alpha/2}\right) = \alpha/2 \right\}$$
$$\hat{\theta}_{1-\alpha/2} = \left\{ \theta : \ Pr\left(\theta_j \ge \hat{\theta}_{1-\alpha/2}\right) = 1 - \alpha/2 \right\} \quad \text{(S6)}$$

S3. Initial model screening

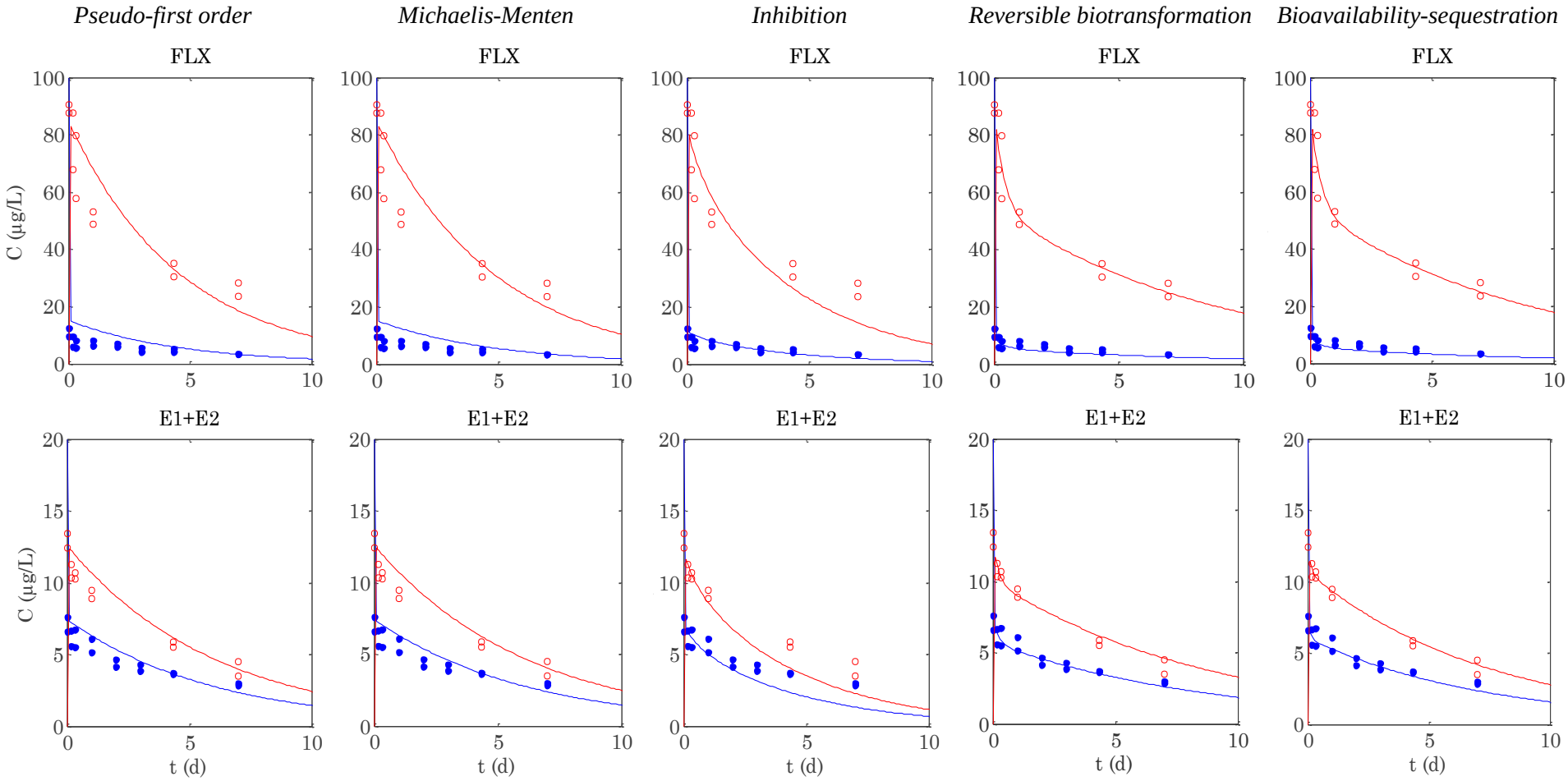

Figure S3a. Experimental (dots) and modelled (lines) concentrations of fluoxetine (FLX) and the hormones (E1+E2) using the five biotransformation hypotheses. Red and blue colours refer to the solid and liquid phase concentrations during methanogenic experiments 1-2.

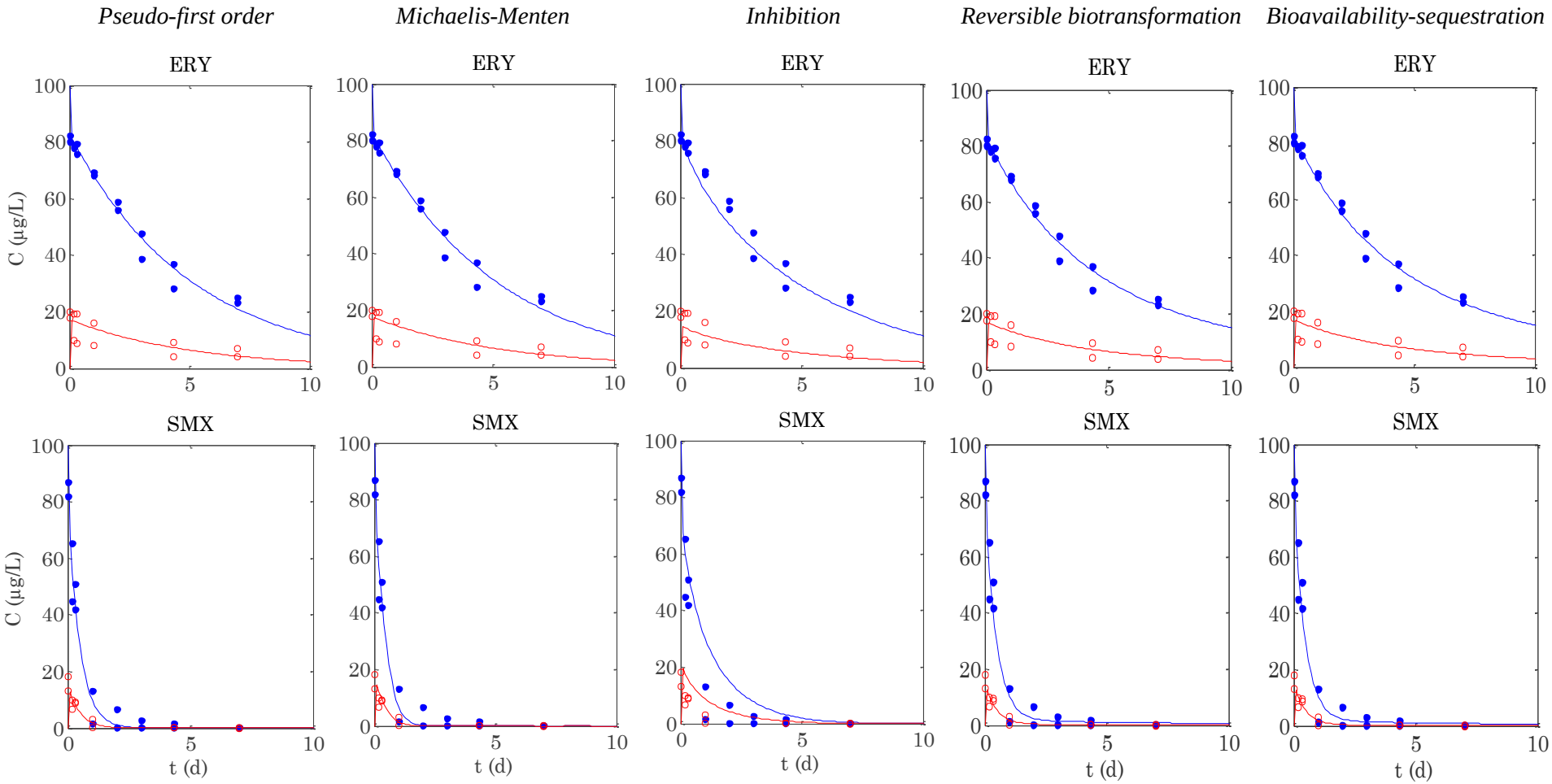

Figure S3b. Experimental (dots) and modelled (lines) concentrations of erythromycin (ERY) and the sulfamethoxazole (SMX) using the five biotransformation hypotheses. Red and blue colours refer to the solid and liquid phase concentrations during methanogenic experiments 1-2.

S4. Model calibration for methanogenesis

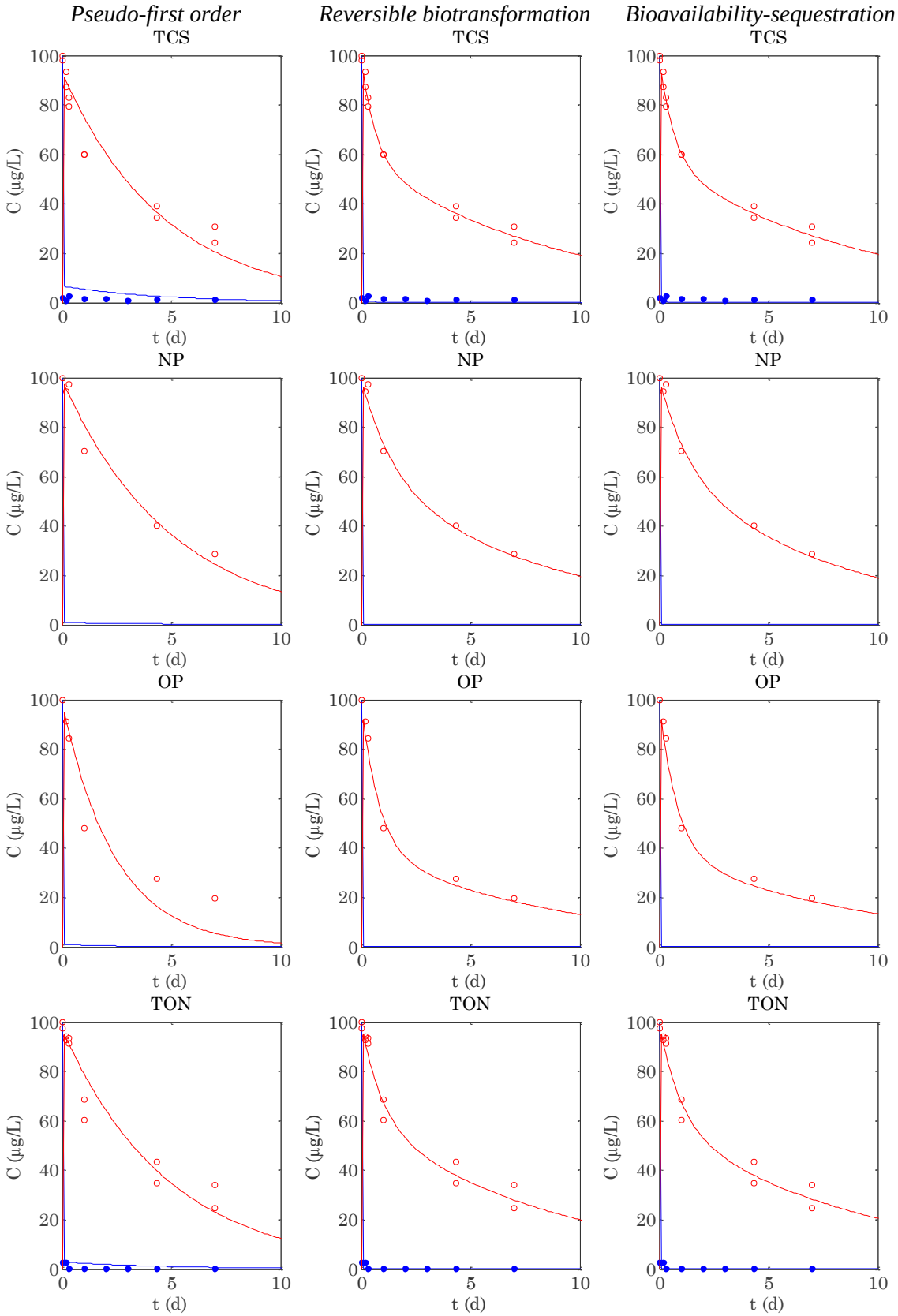

Figure S4a. Experimental (dots) and modelled (lines) concentrations of TCS, NP, OP, and TON (Group 1) during methanogenic experiments 1-2. Red and blue colors refer to the concentration in the solid and liquid phase, respectively.

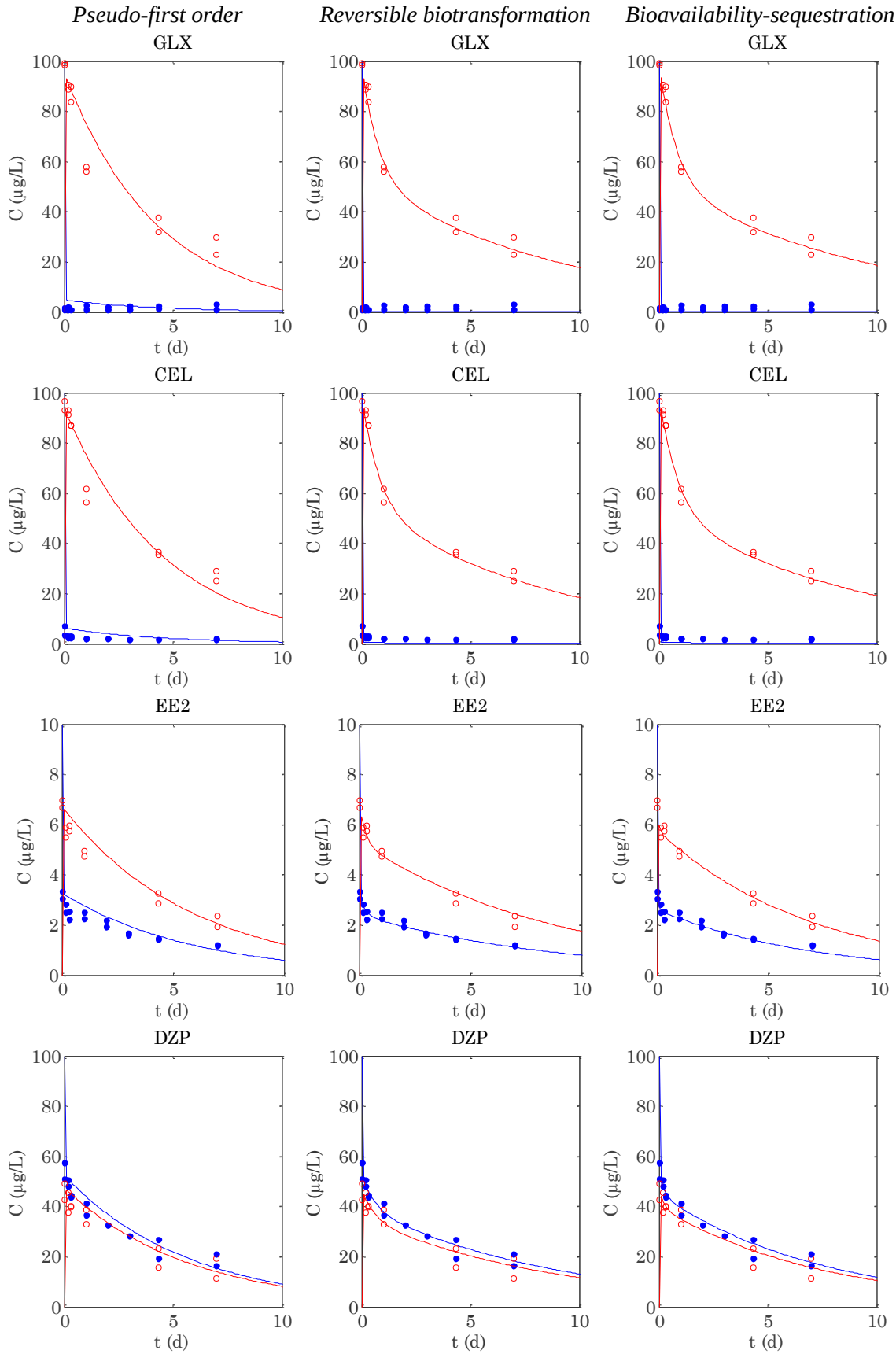

Figure S4b. Experimental (dots) and modelled (lines) concentrations of GLX (Group 1), CEL (Group 1), EE2 (Group 2), and DZP (Group 2) during methanogenic experiments 1-2. Red and blue colors refer to the concentration in the solid and liquid phase, respectively.

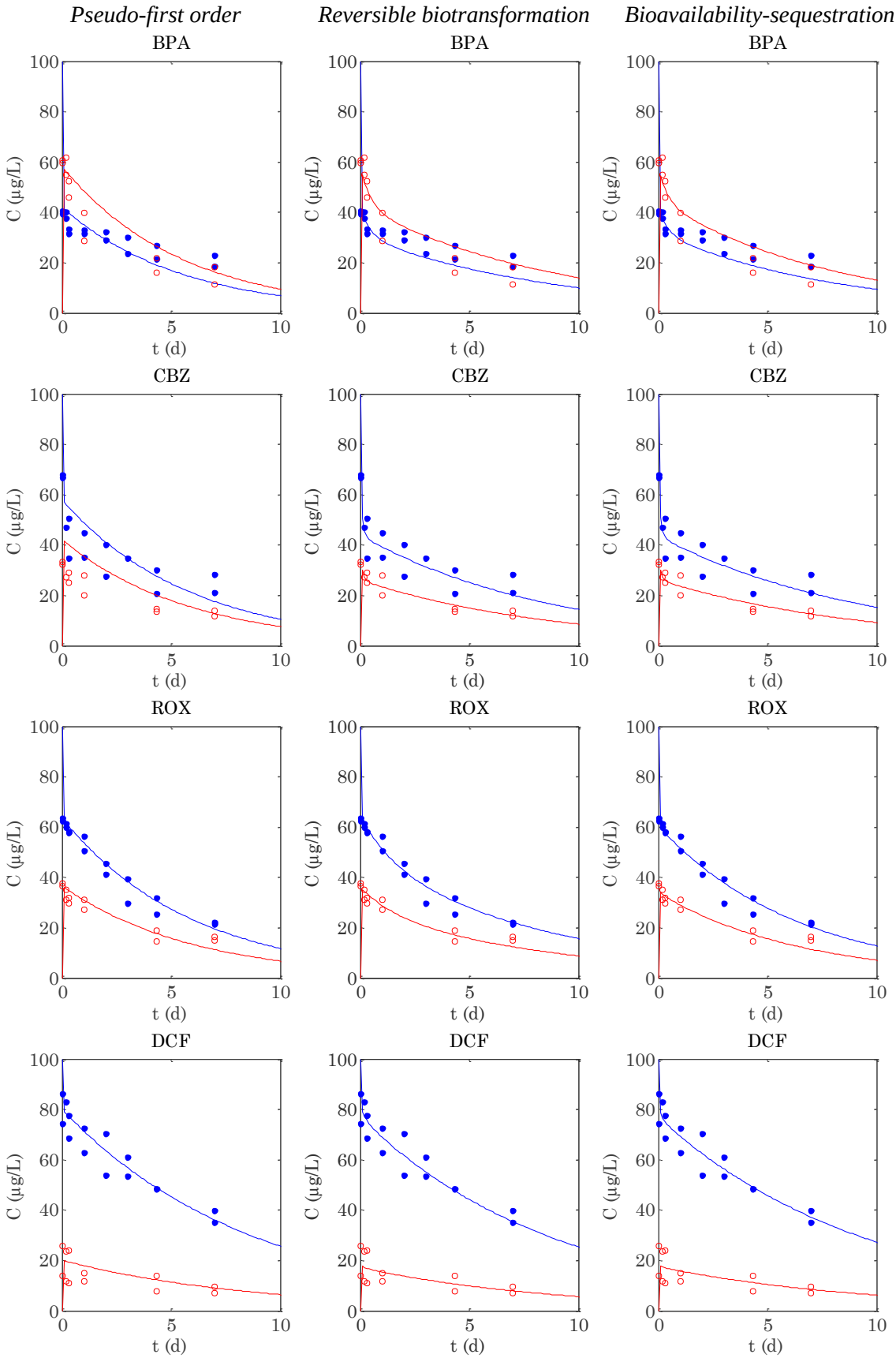

Figure S4c. Experimental (dots) and modelled (lines) concentrations of BPA (Group 2), CBZ (Group 3), ROX (Group 3), and DCF (Group 3) during methanogenic experiments 1-2. Red and blue colors refer to the concentration in the solid and liquid phase, respectively.

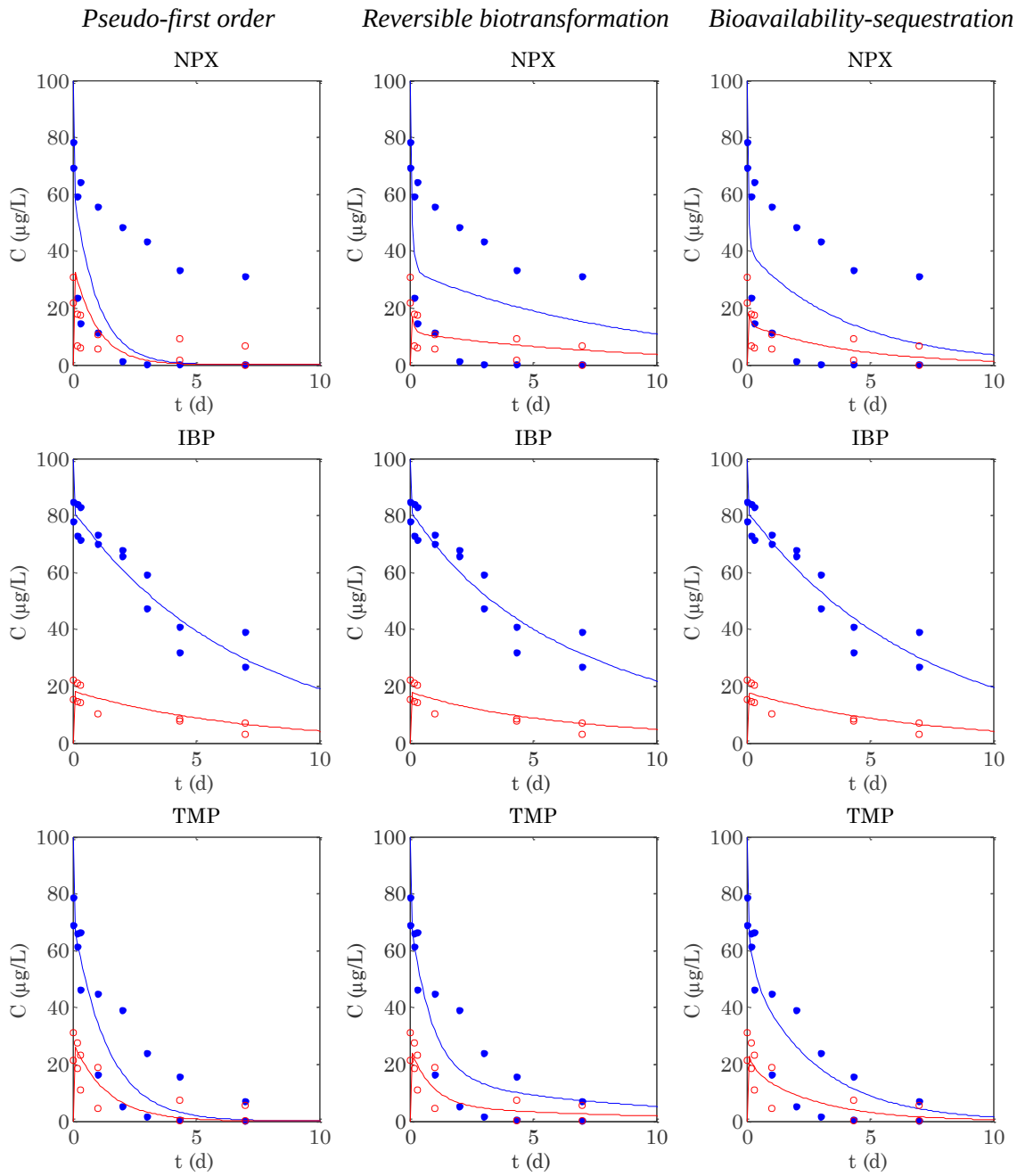

Figure S4d. Experimental (dots) and modelled (lines) concentrations of NPX (Group 3), IBP (Group 3), and TMP (Group 4) during methanogenic experiments 1-2. Red and blue colors refer to the concentration in the solid and liquid phase, respectively.

S5. Model validation for methanogenesis

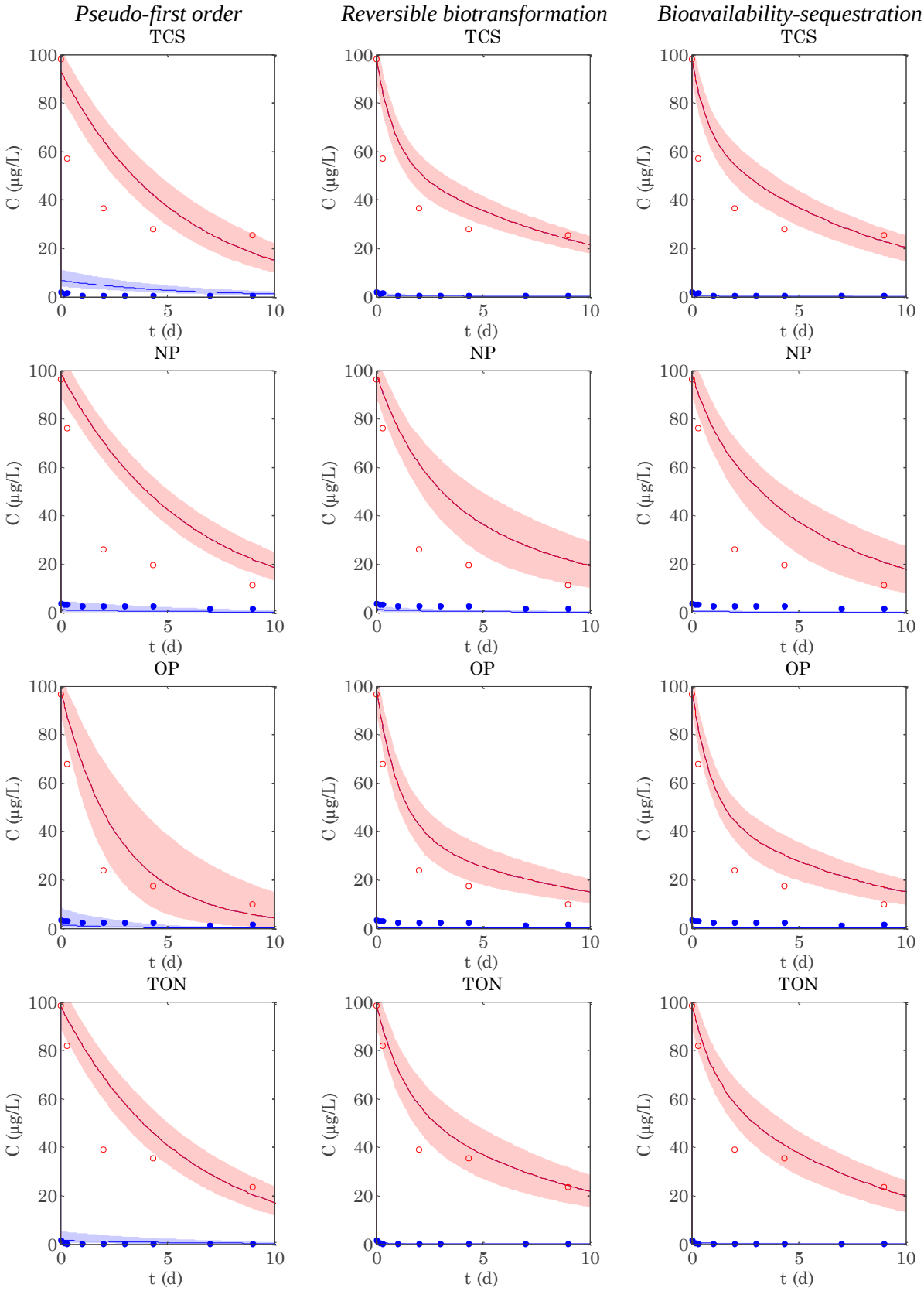

Figure S5a. Model comparison during parameter validation (Experiment 3) of TCS, NP, OP, and TON (Group 1). The blue (liquid) and red (solid) areas represent the fate of the OMP with a 95% of confidence.

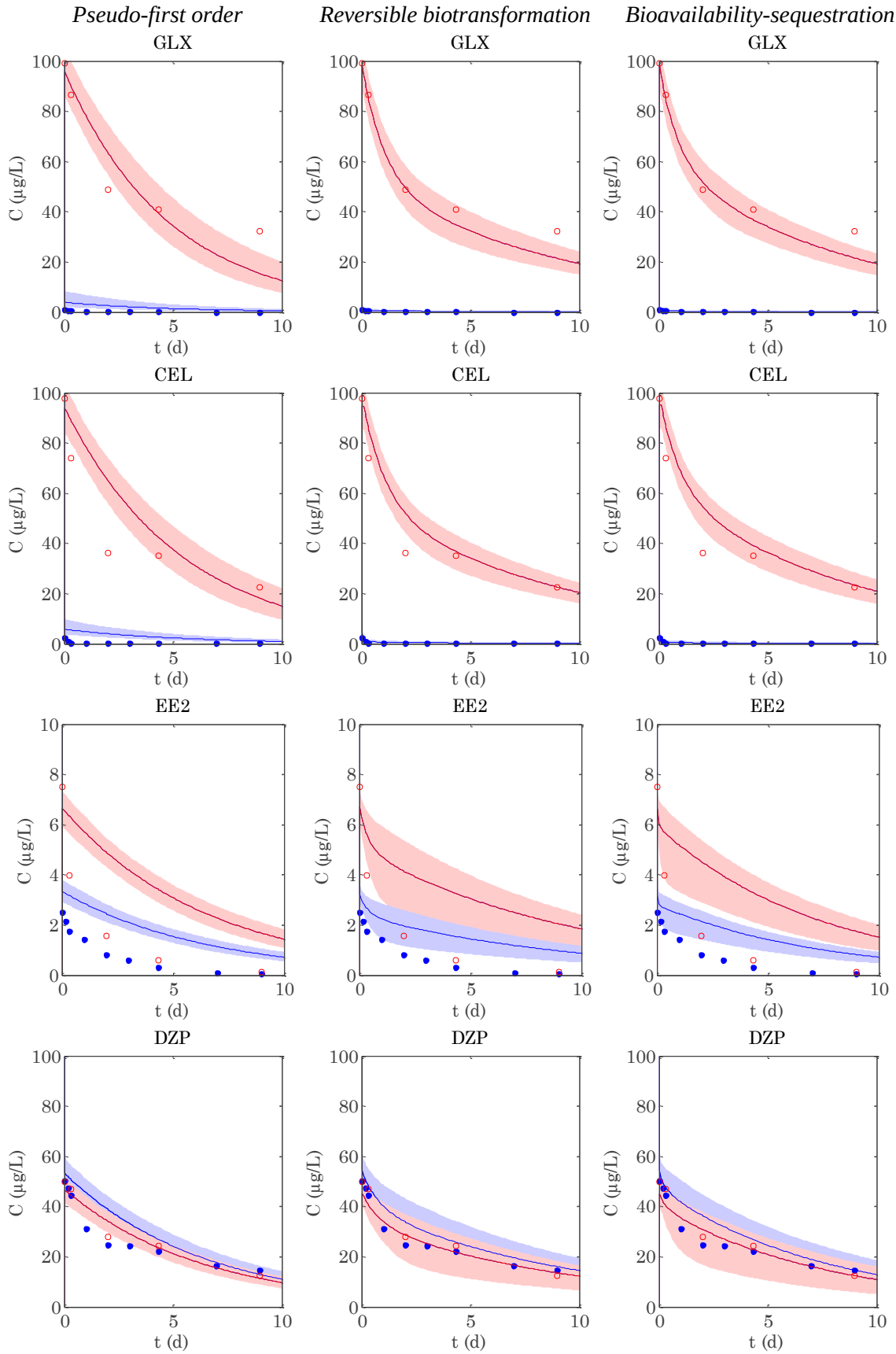

Figure S5b. Model comparison during parameter validation (Experiment 3) of GLX (Group 1), CEL (Group 1), EE2 (Group 2), and DZP (Group 2). The blue (liquid) and red (solid) areas represent the fate of the OMP with a 95% of confidence.

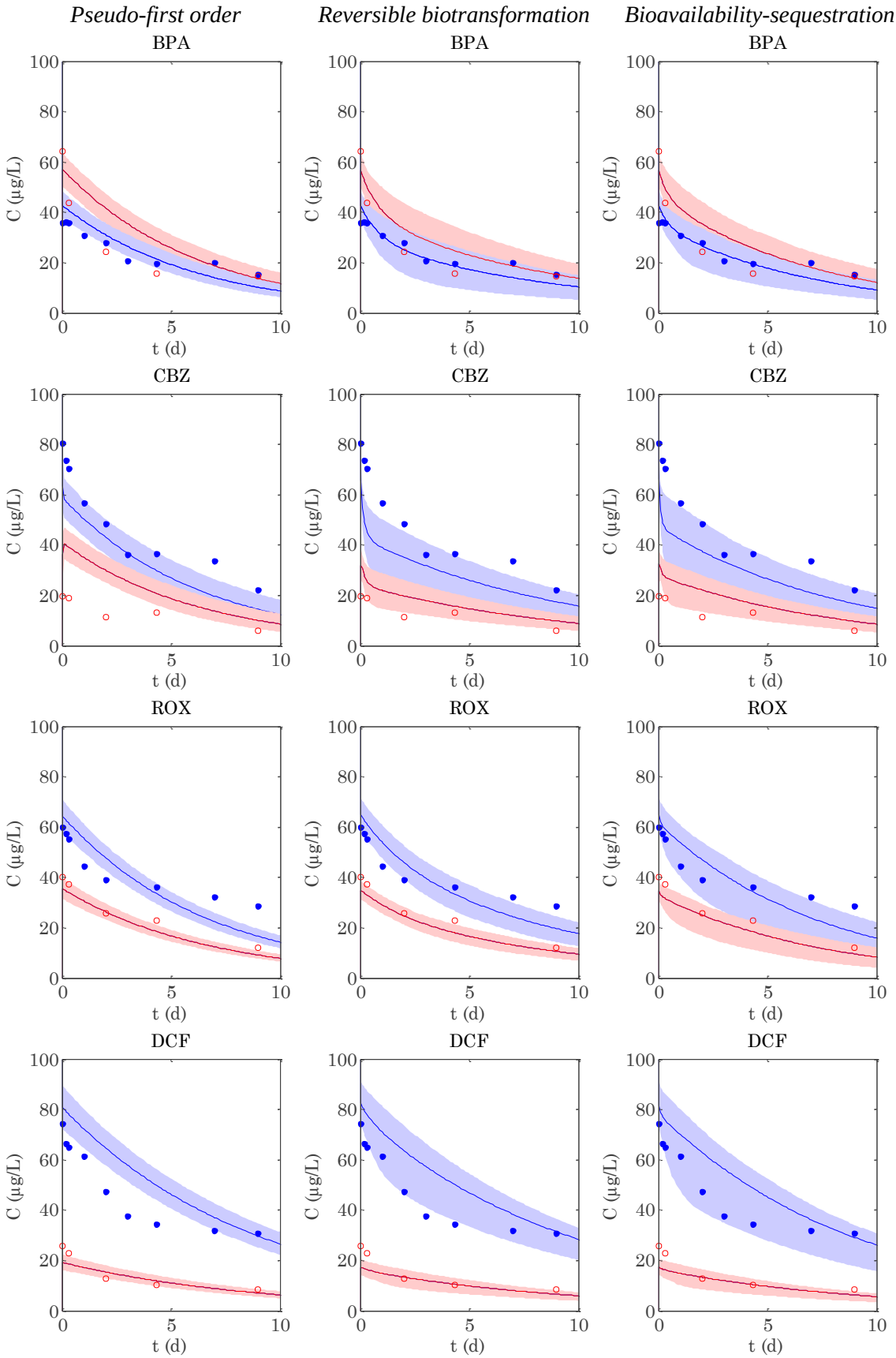

Figure S5c. Model comparison during parameter validation (Experiment 3) of BPA (Group 2), CBZ (Group 3), ROX (Group 3), and DCF (Group 3). The blue (liquid) and red (solid) areas represent the fate of the OMP with a 95% of confidence.

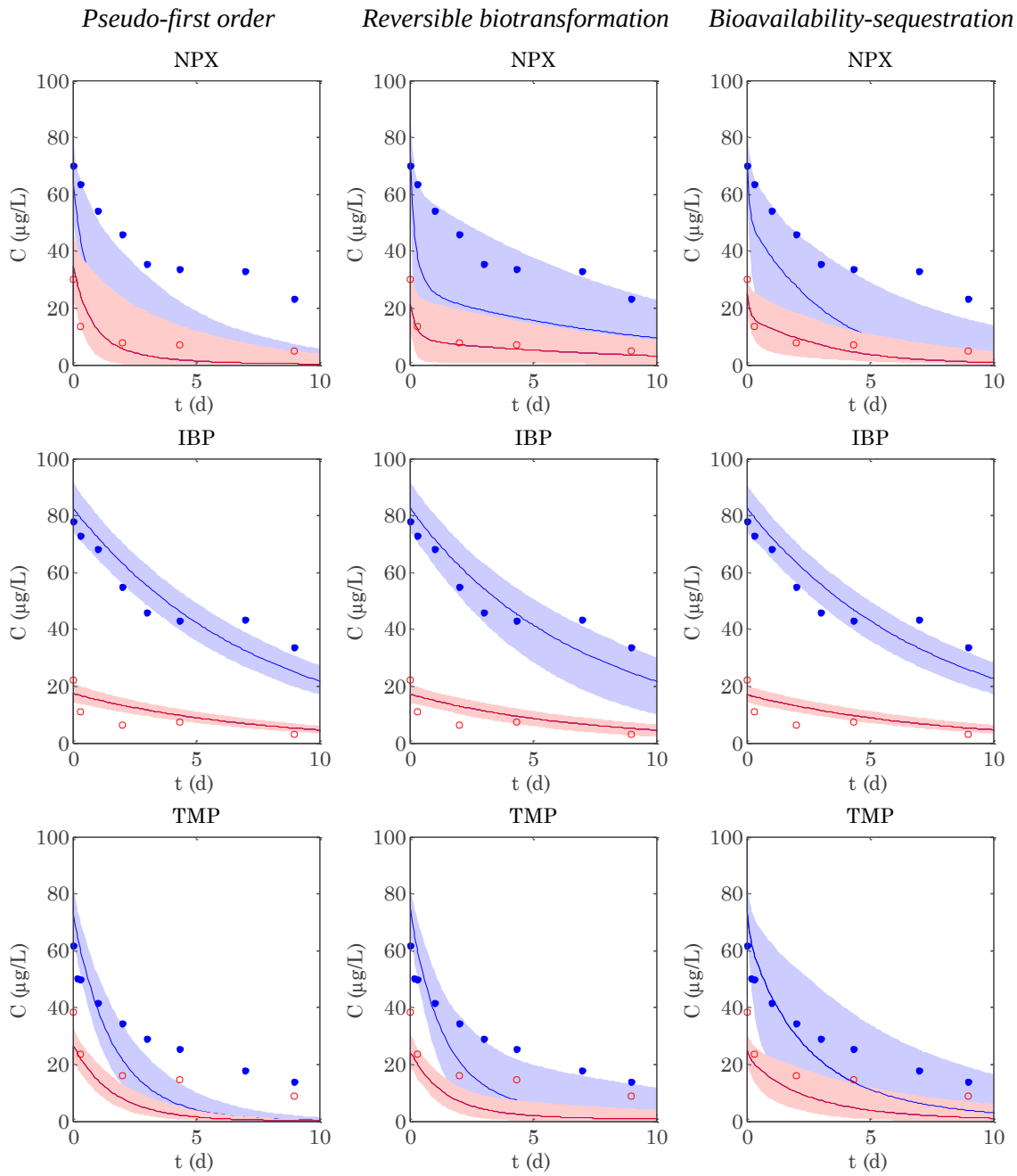

Figure S5d. Model comparison during parameter validation (Experiment 3) of NPX (Group 3), IBP (Group 3), and TMP (Group 4). The blue (liquid) and red (solid) areas represent the fate of the OMP with a 95% of confidence.

S6. Mechanistic model verification in the overall anaerobic process

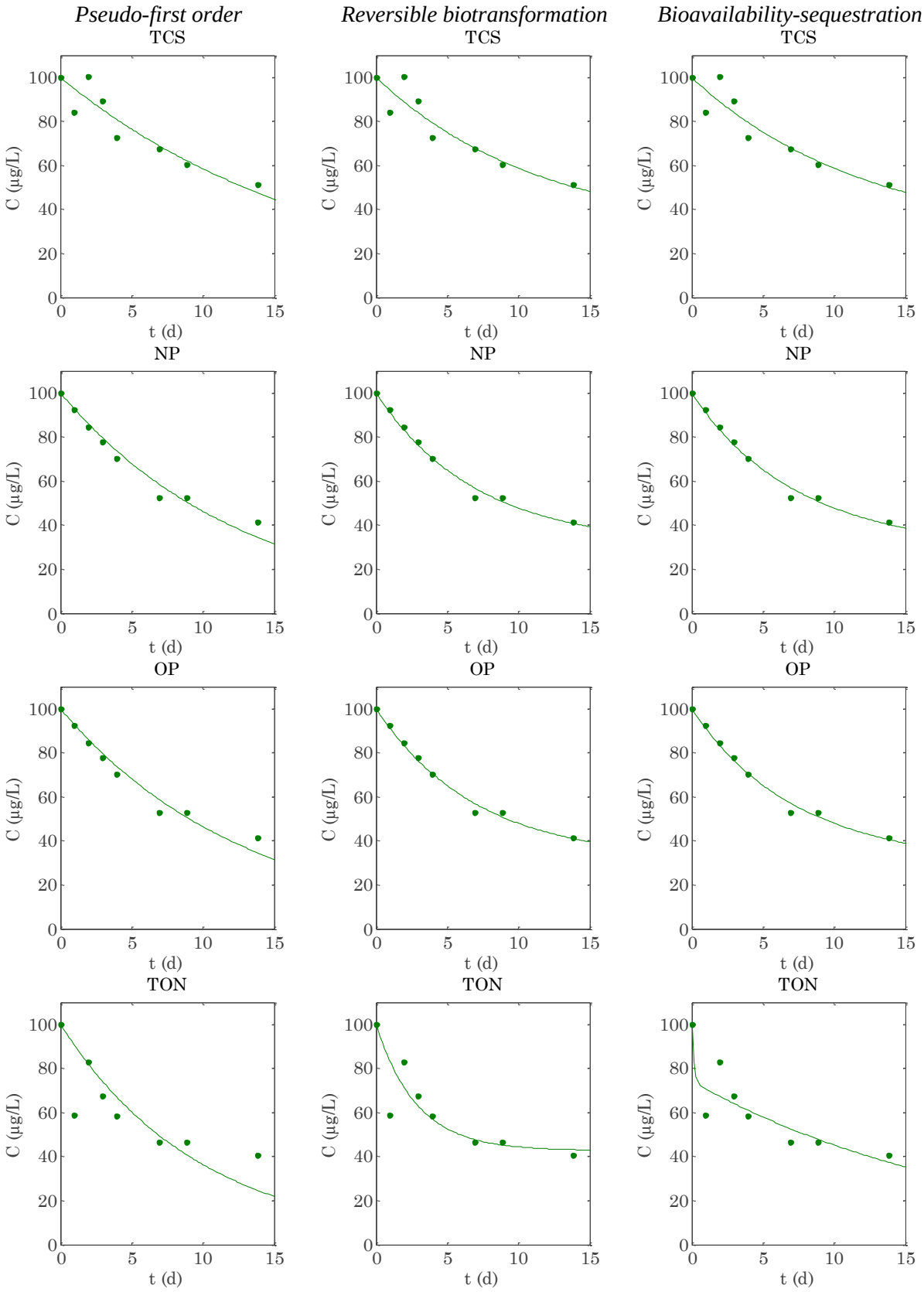

Figure S6a. Model comparison to fit the experimental total concentrations (dots) of TCS, NP, OP, and TON (Group 1) obtained during batch anaerobic digestion (Experiment 4).

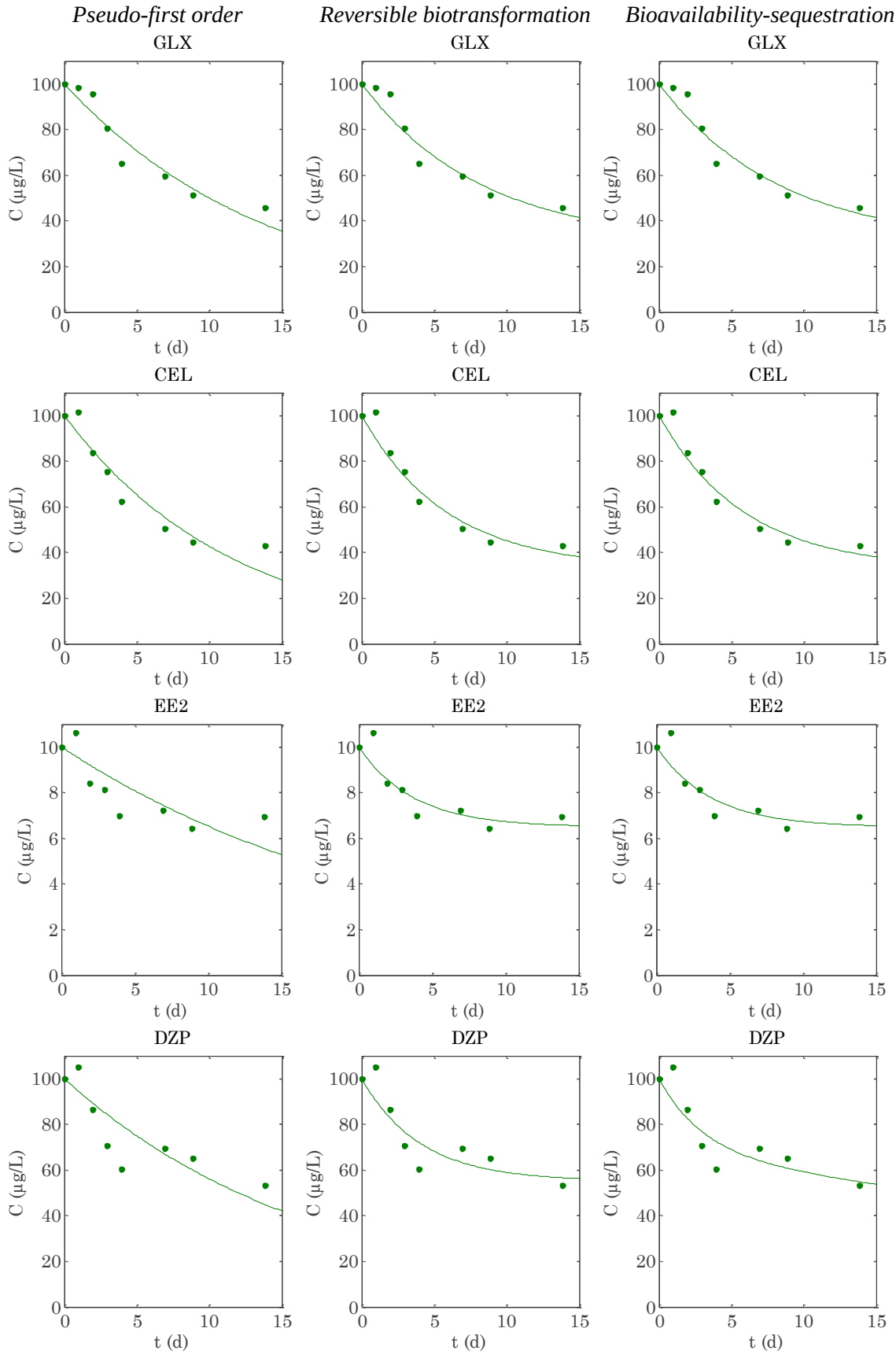

Figure S6b. Model comparison to fit the experimental total concentrations (dots) of GLX (Group 1), CEL (Group 1), EE2 (Group 2), and DZP (Group 2) obtained during batch anaerobic digestion (Experiment 4).

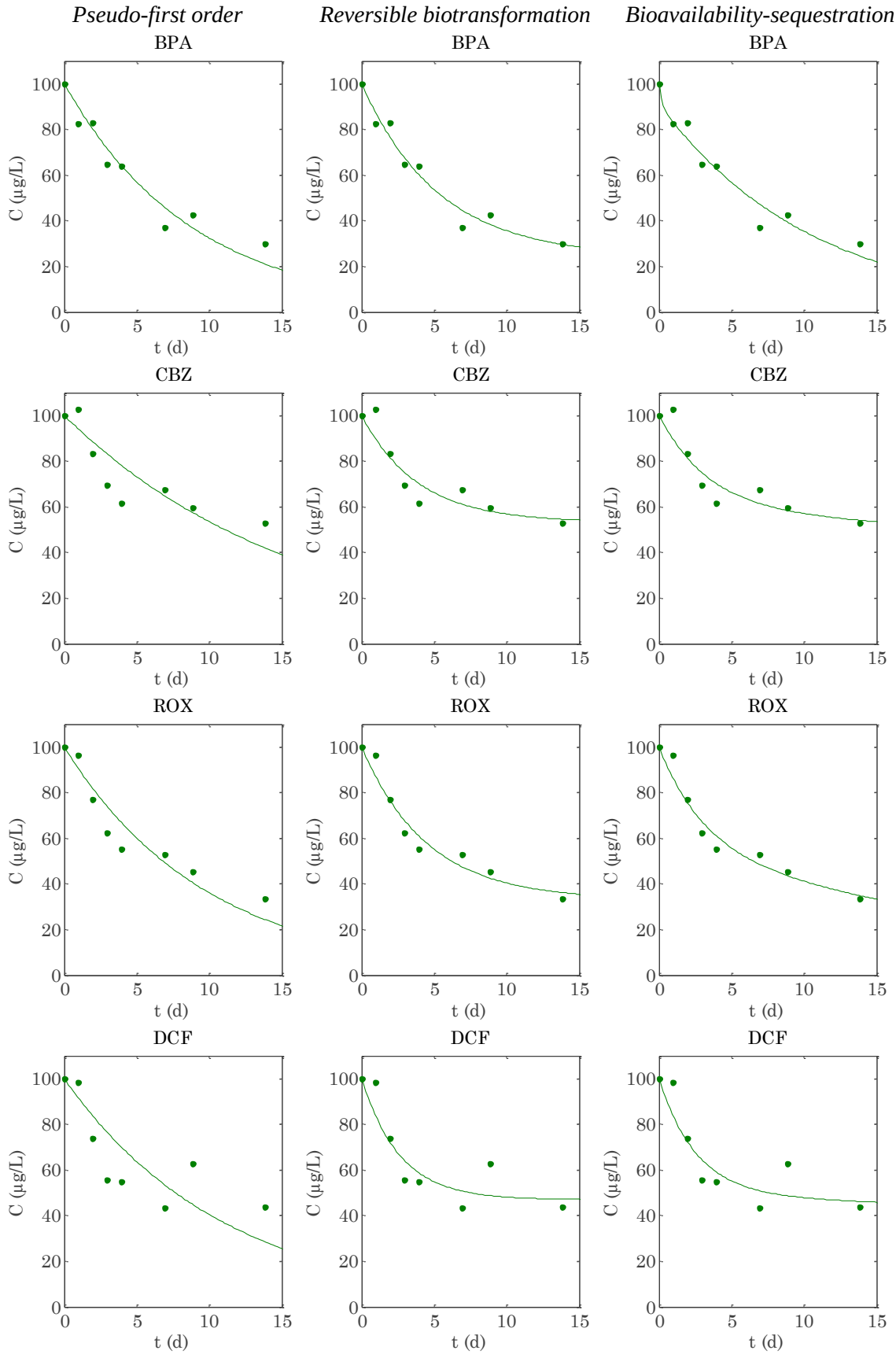

Figure S6c. Model comparison to fit the experimental total concentrations (dots) of BPA (Group 2), CBZ (Group 2), ROX (Group 3), and DCF (Group 3) obtained during batch anaerobic digestion (Experiment 4).

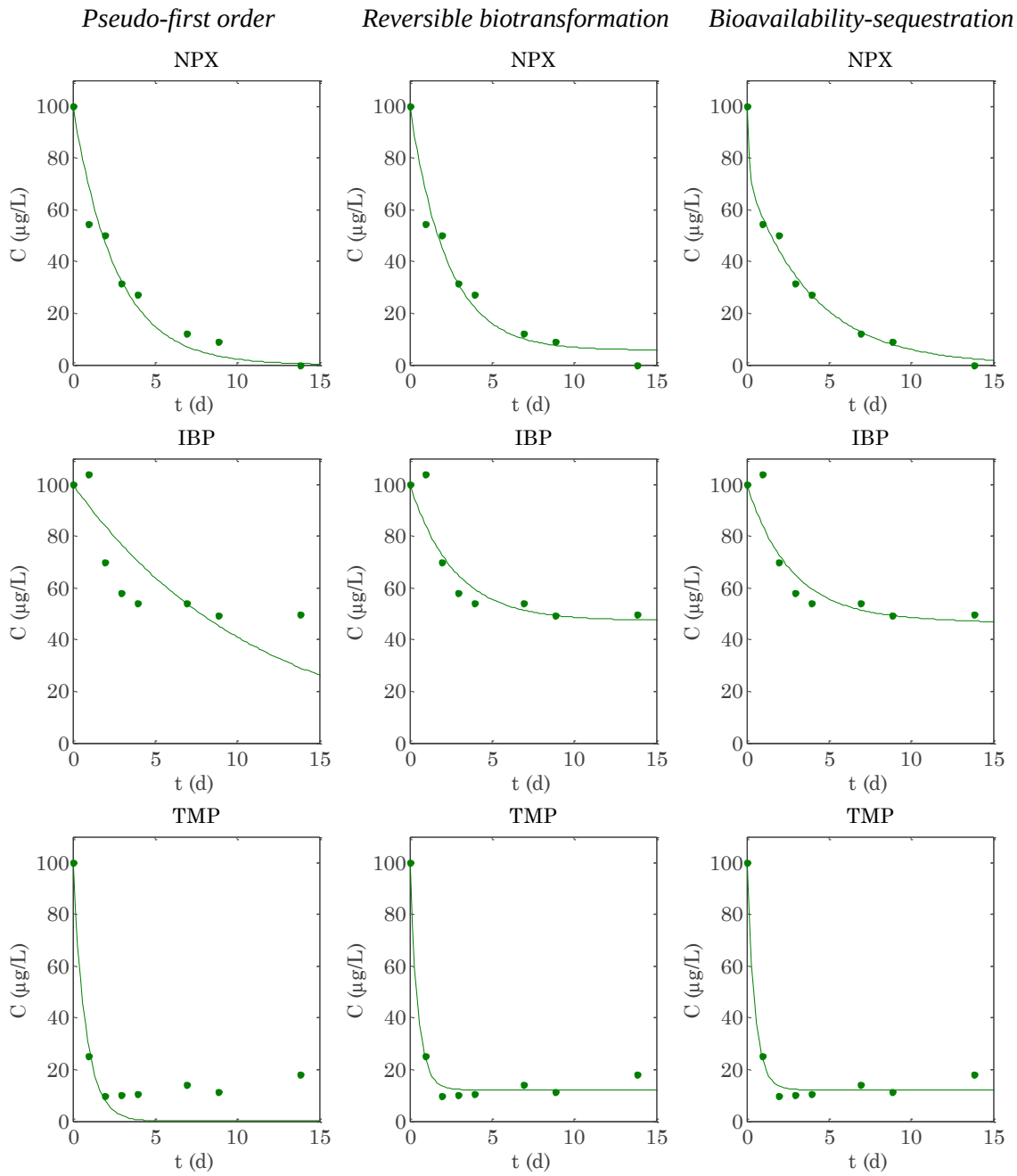

Figure S6d. Model comparison to fit the experimental total concentrations (dots) of NPX (Group 3), IBP (Group 3), and TMP (Group 4) obtained during batch anaerobic digestion (Experiment 4).

S7. Comparison of the RMSE in all of the modelling experiments

Table S7. Root mean square errors (RMSE, µg/L) obtained during calibration and validation with methanogenic data and during calibration with anaerobic digestion data for the pseudo-first order, the reversible biotransformation and the sequestration models.

		Methanogenic calibration (Exp. 1-2)			*Methanogenic validation (Exp. 3)*			*Anaerobic calibration (Exp. 4)*		
	OMPs	Pseudo-first order	Reversible biotransf.	Sequestration	Pseudo-first order	Reversible biotransf.	Sequestration	Pseudo-first order	Reversible biotransf.	Sequestration
Group 1	FLX	8.06	4.29	4.30	9.49	4.56	5.16	9.80	3.45	3.36
	TCS	6.62	2.04	2.01	12.4	8.73	9.39	6.43	6.29	6.29
	NP	5.20	3.59	3.63	15.2	12.1	12.1	3.65	1.83	1.88
	OP	10.3	3.81	3.80	8.55	7.36	7.88	3.67	1.82	1.82
	TON	5.74	3.51	3.46	9.09	5.39	5.74	13.4	9.83	7.53
	GLX	6.04	2.62	2.56	6.47	3.43	3.27	5.95	5.38	5.38
	CEL	5.45	2.51	2.43	9.54	5.44	5.96	6.69	5.06	5.00
Group 2	E1+E2*	0.974	0.505	0.471	1.68	0.952	1.04	1.50	1.18	1.19
	EE2*	0.476	0.226	0.192	1.74	1.45	1.46	0.903	0.604	0.604
	DZP	4.48	3.07	3.07	6.80	4.89	5.73	9.94	7.73	7.73
	BPA	6.43	5.36	5.35	7.91	5.27	5.67	6.11	4.55	5.18
	CBZ	8.97	4.27	4.21	12.6	13.2	12.8	9.24	6.35	6.34
Group 3	ERY	3.49	3.39	3.47	5.65	5.88	5.11	4.02	4.03	4.03
	ROX	3.13	2.66	2.78	6.72	5.85	5.95	7.43	4.90	4.80
	DCF	5.23	4.95	4.91	10.5	10.6	9.3	13.5	8.38	8.47
	NPX	19.3	15.1	15.2	21.8	15.9	13.9	6.23	5.64	2.69
	IBP	4.93	4.98	4.92	6.60	6.43	6.54	13.3	7.83	7.86
G. 4	SMX	4.24	4.21	4.23	12.4	12.4	11.3	0.70	0.70	0.54
	TMP	8.94	8.52	8.93	12.4	12.4	9.25	10.1	2.82	2.75

*The RMSE values of E1+E2 and EE2 are lower than the other OMPs because their initial concentration was 20 µg/L instead of 100 µg/L.