Modelling Emerging Pollutants in Wastewater Treatment: A Case Study using the Pharmaceutical $17\alpha-$ethinylestradiol

Edward Acheampong[a,b,e,*], Ian L. Dryden[a], Jonathan A. D. Wattis[a,*], Jamie Twycross[c], Mark D. Scrimshaw[d], Rachel L. Gomes[b,*]

[a] *School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*
[b] *Food, Water, Waste Research Group, Faculty of Engineering, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*
[c] *School of Computer Science, University of Nottingham, Jubilee Campus, Nottingham, NG8 1BB, UK*
[d] *Department of Life Sciences, College of Health and Life Sciences, Brunel University, Uxbridge UB8 3PH, UK*
[e] *Department of Mathematical Sciences, Methodist University College Ghana, Box DC 940 Dansoman-Accra, Ghana*

Abstract

Mathematical modelling can play a key role in understanding as well as quantifying uncertainties surrounding the presence and fate of emerging pollutants in wastewater treatment processes (WWTPs). This paper presents for the first time a simplified emerging pollutant pathway in the WWTP that incorporates two potential pathways to sequestration. It develops deterministic and stochastic ordinary differential equations to gain insight into the fate and behaviour of a case study pharmaceutical, with particular focus on sorption to the solid phase, as well as the nature of the experimentally measured solid parent compound. Statistical estimation and inferential procedures are developed and via a proof-of-concept examination, the study explores the transformation pathways of the bioactive chemicals (BACs) in

*Corresponding author
Email addresses:
`Edward.Acheampong@nottingham.ac.uk/EAcheampong@mucg.edu.gh` (Edward Acheampong), `Jonathan.Wattis@nottingham.ac.uk` (Jonathan A. D. Wattis), `Rachel.Gomes@nottingham.ac.uk` (Rachel L. Gomes)

the bioreactor, which is the heart of the WWTP. With a focus on the case study pharmaceutical $17\alpha-$ethinlyestradiol (EE2), the simulation results show good agreement with the EE2 data. In addition, the results suggest that the experimentally measured solid EE2-parent concentration is very similar to the model-based sequestered EE2-parent concentration.

Keywords: Differential equations, Pharmaceutical pollutants, Sequestration, Transformation pathways, Wastewater treatment.

---

1. Introduction

Maintaining water quality is a major global challenge due to the increasing use of pharmaceutical drugs and related bioactive chemicals (BACs). Since all water is reused, the urban water cycle requires protection via treatment to ensure that water is of appropriate quality. However, wastewater treatment processes (WWTPs) were never designed to remove emerging pollutants (such as the BACs) from wastewater (Margot et al., 2015; Salgado et al., 2012). Therefore, in order to improve BAC removal it is important to understand the transformation pathways of BACs in the WWTP.

The presence of BACs (e.g., pharmaceutical, steroids) observed in wastewater and the aquatic environment even at low concentrations (ngL^{-1}) can cause adverse effects in aquatic organisms (Verlicchi et al., 2012). Hence, increasingly stringent legislation such as the EU Water Framework Directive (WFD) 2015/495 requires those responsible for wastewater treatment to monitor the presence of these pollutants with the future aim of improving the quality of the wastewater effluent released to surface waters (Carvalho et al., 2015; Barbosa et al., 2016). This has resulted in a significant number of investigations into the removal of BACs in the WWTP. Recent studies include mathematical models that take into consideration the influence of free (or parent) compounds as well as retransformable products such as conjugated compounds (Xu et al., 2016), where conjugated compounds are Phase II metabolites formed as a result of drug transformation processes (Coleman, 2010).

Sorption to activated sludge plays a role in the fate of BACs in WWTPs and is dependent on the physiochemical properties of the BAC. Sorption can be a precursor to biotransformation/biodegradation. Several experimental sorption studies have found the occurrence of sequestration, where pollutants associated with the solid phase (e.g. mixed liquor suspended solids (MLSS), sludge or sediment) is not readily reversible (Yi and Harper, 2007; Barret et al., 2011; Sittig et al., 2012). However, there is limited mathematical modelling insight into this phenomenon and no studies to the best of the authors knowledge show whether the BACs concentration on the solid phase is on the surface of the activated sludge (adsorption), within (absorption) or a combination of both (Gomes et al., 2009; Banihashemi and Droste, 2014). Distinguishing between adsorption and absorption is vital since the absorbed BACs are not readily available for microbial biodegradation (Banihashemi and Droste, 2014). That is, the ratio of BACs adsorbed to absorbed has implications in terms of the distribution from liquid phase to solid phase and vice versa (Yi and Harper, 2007; Barret et al., 2011). This has not been the focus of many investigations but as Barret et al. (2011) suggested, the quantification of sorption of BACs with activated sludge will aid the prediction of the fate of BACs and the associated risk to the environment.

The majority of mathematical modelling research carried out in the field of WWTP (lab-, pilot- or full-scale) employs deterministic ODE modelling based on the assumption that all the biological, chemical and physical laws governing the system can be understood (Gernaey et al., 2004). For example, mathematical models describing the fate of both traditional and emerging pollutants in the WWTP have been placed under the general framework of the activated sludge model for Xenobiotics (ASM-X) (Plósz et al., 2010; Snip et al., 2014; Polesel et al., 2015). Nevertheless, modelling of WWTP is adversely influenced by imperfect knowledge of biochemical processes, operational parameters, environmental factors and input loadings. These variabilities have both temporal and spatial effects in the removal of emerging pollutants from treated wastewater. Liu et al. (2015) pointed out that while lab-scale batch experiments have a role, they do not often reflect full-scale wastewater treatment or integrated wastewater treatment observations

because wastewater is stochastic in nature. Stochastic models allow the estimation of model parameters and the understanding of the uncertainties in pollutant concentrations.

EE2 is a pharmaceutical and synthetic estrogen used in birth control pills (Laurenson et al., 2014). A review conducted by Ting and Praveena (2017) reveals that the excretion rate for women who used such medication is $35\mu\text{gday}^{-1}$ and that 80% of EE2 taken into the body is excreted as unmetabolized conjugates with 30% from faeces and 22-50% from urine. Due to the potential adverse endocrine disrupting effects associated with even small (ngL^{-1}) concentrations in the aquatic environment, it is reported in the Watch List of the EU WFD 2015/495 (Carvalho et al., 2015; Barbosa et al., 2016). Thus, EE2 is considered as the case study BAC in this paper and the pathway of EE2 in WWTPs proposed by Joss et al. (2006), which is widely applied to model the fate of most emerging pollutants, is extended by introducing the concept of sequestration to investigate the relationship and nature of BACs and the solid phase. The impetus for this research is the fact that most work on sorption used an experimental (Yi et al., 2011) rather than a mathematical modelling approach. It should be noted that although Polesel et al. (2015) and Snip et al. (2014) considered the formation of sequestered products in their full-scale wastewater treatment modelling for other pharmaceuticals, they assumed that the ratio of sequestered concentration to pre-clarified influent concentration of the free parent compound is constant and as such, no kinetic rate parameter is accounted for the sequestration mechanism.

Mathematical modelling saves time and the cost of repeatedly designing new experiments by offering the opportunity to alternatively deepen knowledge of complex mechanisms via computer simulation evaluations. For this reason, a mathematical modelling approach using experimental batch data taken from an activated system (Gomes et al., 2009) is used with the aim of extracting the maximum information on the mechanisms involved in EE2 treatment by fitting a range of advanced mathematical models, including both deterministic, stochastic and combined approaches, to a dataset. This study is concerned with four forms of EE2, which are: the parent chemical

in the liquid phase (free); adsorbed onto activated sludge (sorbed), absorbed into activated sludge (sequestered) and those retransformable forms of the parent chemical (e.g., excreted Phase I and II metabolites) as summarised in Table 1 of Section 3. An experimental dataset from Gomes et al. (2009) is selected for the study since it captures the changes in the concentrations of EE2 states over time. Data of such a nature are limited in the literature as more usually data involve monitoring influent and effluent only (Ting and Praveena, 2017; Yu et al., 2019).

2. Experimental data

The data are obtained from batch scale fate studies using activated sludge from a pilot plant following a Husmann design (Gomes et al., 2009). A detailed description of the Husmann pilot plant used by Gomes et al. (2009) follows the OECD guidelines for testing chemicals in an activated sludge process (OECD, 2001). The dataset contains measurements of concentrations at eight different time points for three EE2 forms: the first form is the aqueous parent EE2, the second is the aqueous conjugated steroid ethinylestradiol 3 - glucuronide (EE2 – 3G) and the third one is the solid parent EE2. The measurements are recorded in hours (h) at 0, 0.0167, 0.5, 1, 2, 4, 8 and 24 hours in triplicate. The mixed liquor suspended solid concentration (C_{MLSS}) reported is 4 gL^{-1} (Gomes et al., 2009). For a detailed description of the experimental setup see Gomes et al. (2009). For the purpose of mathematical modelling, Y is used to denote the measured EE2 concentrations, where Y_1 is the liquid form of parent EE2, Y_2 is the EE2-3G and Y_3 is the solid form of parent EE2.

3. Dynamical model

3.1. Model description

A computational model was developed to describe the fate of BAC(s) in the unit treatment technology utilising lab-scale data. In experimental studies, the activated sludge is usually spiked with the BAC(s) of interest,

which are classified by four state variables, namely, the liquid form of parent compound, C_{free}, liquid form of retransformable compounds, C_{con}, sorbed parent compounds, C_{sor}, and sequestered parent compounds, C_{seq} (Polesel et al., 2015). The activated sludge bioreactor is assumed to contain a constant biomass (MLSS) denoted by $C_{\mathrm{MLSS}} = 4\,\mathrm{gL}^{-1}$ (Gomes et al., 2009) since the BAC concentrations are in the range of mgL^{-1} to ngL^{-1}, the biomass growth is negligible during the consumption of the BACs and any effects of the BAC on the actual biomass activity are assumed to be negligible (Snip et al., 2014). The following additional assumptions are made:

1. the reactor, of volume V, is well-mixed and operates on a continuously stirred tank reactor (CSTR),

2. the dissolved oxygen concentration in the reactor is constant,

3. the retransformable compounds (C_{con}) are not further biotransformed, but may be deconjugated back to the liquid form of parent compounds (C_{free}) (Joss et al., 2006; Gomes et al., 2009; Kumar et al., 2012; Snip et al., 2014),

4. the dissolved forms of parent compounds (C_{free}) in liquid are in a reversible reaction with the solid (sorbed) fraction (C_{sor}) and are further biodegraded via microbial activities as well as irreversibly absorbed, i.e., sequestered to C_{seq}.

5. C_{MLSS} is assumed constant.

Figure 1 shows the simplified pathways diagram, which is modelled using first-order and pseudo-first order kinetics. Table 1 gives a summary of the notation, description of the model parameters and their corresponding units.

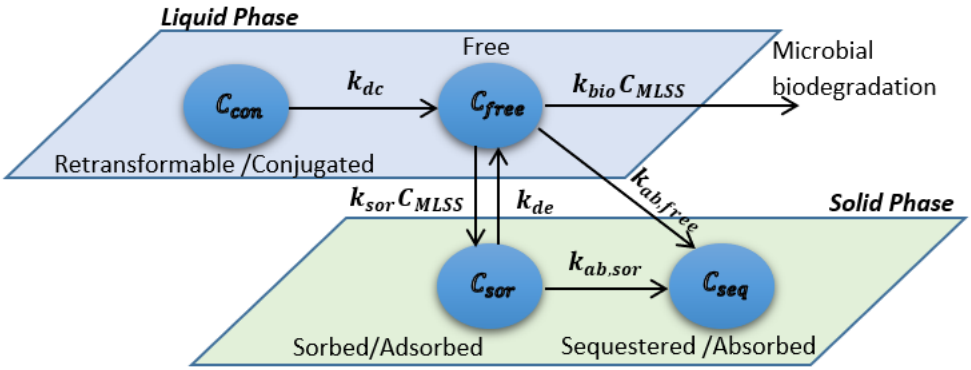


Fig. 1. Diagram of the simplified pathways for emerging pollutants in the bioreactor of the wastewater treatment processes. See Table 1 for an explanation of parameters and concentration variables.

Table 1
Model parameters, state variables and other model notations

Variable	Description	Measured data	Unit
C_{con}	EE2-conjugated: Liquid retransformable compound concentration	Y_1	$\mathrm{ngL^{-1}}$
C_{free}	EE2-free: Liquid parent compound concentration	Y_2	$\mathrm{ngL^{-1}}$
C_{sor}	EE2-sorbed: Adsorbed solid parent compound concentration		$\mathrm{ngL^{-1}}$
		Y_3	
C_{seq}	EE2-sequestered: Absorbed solid parent compound concentration		$\mathrm{ngL^{-1}}$
Parameter($\boldsymbol{\theta}$)	Description		Unit
C_{MLSS}	Mixed liquor suspended solid concentration (assumed fixed)		$\mathrm{gL^{-1}}$
k_{dc}	Deconjugation rate coefficient for C_{con}		$\mathrm{h^{-1}}$
k_{sor}	Sorption rate coefficient for C_{sor}		$\mathrm{Lg^{-1}h^{-1}}$
k_{de}	De-sorption rate coefficient for C_{sor}		$\mathrm{h^{-1}}$
k_{bio}	Biotransformation rate coefficient for C_{free}		$\mathrm{Lg^{-1}h^{-1}}$
$k_{ab,free}$	Absorption rate coefficient for C_{free}		$\mathrm{h^{-1}}$
$k_{ab,sor}$	Absorption rate coefficient for C_{sor}		$\mathrm{h^{-1}}$
f_{sor}	Fraction of sorbed concentration C_{sor}		–
Notation	Description		Unit
k	Number of state variables		–
k'	Number of experimental measured state variables		–
p	Number of model parameters		–
Q	Number of repeated experiments		–
N	Number of experimental time points		–

Note that EE2-sorbed/EE2-sequestered is EE2-parent concentration in the solid phase (EE2-solid) and EE2-free is liquid EE2-parent.

3.2. Bioreactor model

3.2.1. Deterministic ODE model

Figure 1 depicts the interaction between $C_{free}, C_{con}, C_{sor}$ and C_{seq}, it can be seen that microbial biodegradation and sorption activities are the two main mechanisms through which the EE2-parent can be removed. Evolution of the four concentration forms of EE2 within the bioreactor can be described

by four coupled nonlinear systems of ODEs given as

$$\begin{aligned}\frac{dC_{free}}{dt} =& k_{dc}\left(\frac{M_{free}}{M_{con}}\right)C_{con} + k_{de}C_{sor} - k_{sor}C_{free}C_{\mathrm{MLSS}} \\ &- (k_{ab,free} + k_{bio}C_{\mathrm{MLSS}})C_{free},\end{aligned} \tag{3.1}$$

$$\frac{dC_{con}}{dt} = -k_{dc}\left(\frac{M_{free}}{M_{con}}\right)C_{con}, \tag{3.2}$$

$$\frac{dC_{sor}}{dt} = k_{sor}C_{free}C_{\mathrm{MLSS}} - k_{de}C_{sor} - k_{ab,sor}C_{sor}, \tag{3.3}$$

$$\frac{dC_{seq}}{dt} = k_{ab,free}C_{free} + k_{ab,sor}C_{sor}. \tag{3.4}$$

With respect to the lab-scale batch experimental set-up, one state variable has non-zero initial concentration, $C_{con}(0) = C_{con\,0} = 2125$ ngL^{-1} (Gomes et al., 2009) and the remaining state variable concentrations are all zero ($C_{free}(0) = 0, C_{sor}(0) = 0$ and $C_{seq}(0) = 0$). Note that M_{free} (g mol^{-1}) and M_{con} (g mol^{-1}) are respectively the molecular masses of C_{free} and C_{con} and the fraction M_{free}/M_{con} corrects for the change from one compound form to another.

The novelty of the proposed model given by Equations (3.1) - (3.4) lies in the modelling of the sequestration dynamics (as stated in assumption 4 of Section 3.1) at the lab-scale bioreactor. As mentioned earlier, previous studies at the full-scale WWTP by Snip et al. (2014) assumed a constant ratio of the influent liquid parent compound concentration (C_{free}) and solid (sequestered) parent compound concentration (C_{seq}). However, in the proposed model, the kinetic rate parameters $k_{ab,free}$ and $k_{ab,sor}$ account for the absorption of both C_{free} and C_{sor} respectively. The ratio of BACs adsorbed to absorbed have implications for the actual fate and the availability to and from the liquid and solid phases (Yi and Harper, 2007). As a result, the quantification of sequestered BACs will offer a useful insight and help reduce the variability surrounding the fate of BACs in wastewater treatment.

To take into consideration the fact that experimental data are likely to have no sequestration or variation in sequestration pathways since experimental data rarely distinguishes between adsorbed and absorbed BACs.

Therefore, modifications to the proposed pathways (Figure 1) are made leading to $\mathcal{M}0$, $\mathcal{M}1$ and $\mathcal{M}2$ as depicted in Figure 2. The pathways $(1), (2), (3)$ and (4) are well established pathways used in modelling the fate of BACs in wastewater treatment (Joss et al., 2006; Plósz et al., 2012b; Polesel et al., 2015; Plósz et al., 2012a) and so is included in all three models. Thus, the three different sets of pathways this study wish to compare are defined as follows:

- $\mathcal{M}0$, where all pathways are included,
- $\mathcal{M}1$, where sequestration via sorbed parent BACs is neglected,
- $\mathcal{M}2$, where sequestration via liquid parent BACs is neglected.

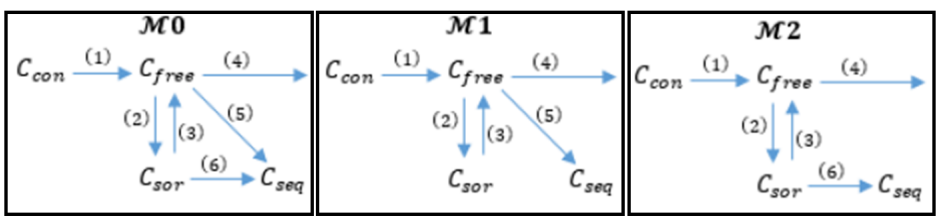


Fig. 2. Proposed pathways of EE2 interactions in wastewater treatments systems. Sequestration formation via pathways (5) and (6) ($\mathcal{M}0$) (**left**), sequestration formation via pathway (5) ($\mathcal{M}1$) (**middle**) and sequestration formation via pathway (6) ($\mathcal{M}2$) (**right**).

The ODE bioreactor model of BACs as given by equations (3.1) - (3.4) and modifications can be written in the general framework given by the following four-dimensional system of ODEs for the underlying state variable, $\boldsymbol{X}_t$, as

$$\frac{d\boldsymbol{X}_t}{dt} = \boldsymbol{f}(\boldsymbol{X}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}_1), \tag{3.5}$$

with $\boldsymbol{X}_0 = [0, C_{con,0}, 0, 0]^T$ as initial state concentrations. Note that, t denote time;

$$\boldsymbol{X}_t = [C_{free}, C_{con}, C_{sor}, C_{seq}]_t^T, \tag{3.6}$$

are the concentrations at time t; T superscript denote matrix transposition;

$$\boldsymbol{\theta}_1 = [k_{dc}, k_d, k_{de}, k_{bio}, k_{ab,free}, k_{ab,sor}], \tag{3.7}$$

is the rate parameter vector; and $\boldsymbol{f}(\boldsymbol{X}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}_1)$ is a nonlinear function with $\boldsymbol{f}(\cdot)$ describing the nonlinear relationship between the underlying state variable, $\boldsymbol{X}_t$ and other input parameters, i.e., $\boldsymbol{u}_t = C_{\text{MLSS}}$ as the input parameter. For simplicity, in the method below the notation $\boldsymbol{X}_t$ and $\boldsymbol{\theta}_1$ are rather used than that of Table 1.

3.2.2. Stochastic ODE model

The state of the biological treatment phase of the WWTP is affected by a number of factors (temperature, pH, product formation, mixing, etc.), which cannot be precisely accounted for by the ODE model (3.5) (Sathyamoorthy et al., 2014). As a result, processes such as conjugation, deconjugation, biotransformation, sorption, desorption, are affected by stochastic fluctuations in the concentrations due to (i) imprecise knowledge of rate coefficients, $\boldsymbol{\theta}$ in (3.5) (e.g., dependency on temperature, pH, etc.) (ii) structural misspecification, $\boldsymbol{f}(\cdot)$ in (3.5) due imperfect knowledge of the biological, chemical and physical processes (e.g., some compounds/products acting as inhibitors to other reactions) and (iii) fluctuations in variable concentrations due to, for example, imperfect mixing as well as varying load entering the treatment process.

To account for the uncertainties and simplifications in the previous model given by equations (3.1) - (3.4), an extension to the SDEs is considered by incorporating multiplicative system noise into the system by equations (3.1) - (3.4) via

$$\begin{aligned} dC_{free} = \Bigg(k_{dc}\left(\frac{M_{free}}{M_{con}}\right) C_{con} + k_{de} C_{sor} - k_{sor} C_{free} C_{\text{MLSS}} - \\ (k_{ab,free} + k_{bio} C_{\text{MLSS}}) C_{free} \Bigg) dt + \sigma_{free} C_{free} d\mathbb{W}_{1t}, \end{aligned} \tag{3.8}$$

$$dC_{con} = \left(- k_{dc}\left(\frac{M_{free}}{M_{con}}\right) C_{con} \right) dt + \sigma_{con} C_{con} d\mathbb{W}_{2t}, \tag{3.9}$$

$$dC_{sor} = \left(k_{sor} C_{free} C_{\text{MLSS}} - k_{de} C_{sor} - k_{ab,sor} C_{sor} \right) dt + \sigma_{sor} C_{sor} d\mathbb{W}_{3t}, \tag{3.10}$$

$$dC_{seq} = \left(k_{ab,free} C_{free} + k_{ab,sor} C_{sor} \right) dt + \sigma_{seq} C_{seq} d\mathbb{W}_{4t}, \tag{3.11}$$

where $\mathbb{W}_t = [\mathbb{W}_1, \mathbb{W}_2, \cdots, \mathbb{W}_d]_t^T$ denotes a standard Wiener process, also known as Brownian motion (Øksendal, 2003). The system given by equations (3.8) - (3.11) can be written in a general framework by the following four-dimensional system of SDEs for the underlying state variable, $\boldsymbol{X}$, as

$$d\boldsymbol{X_t} = \boldsymbol{f}(\boldsymbol{X}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}_1)dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}_1)d\mathbb{W}_t, \tag{3.12}$$

where $\boldsymbol{f}(\boldsymbol{X}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}_1)dt$ denotes the drift term with $\boldsymbol{f}(\cdot)$; representing the deterministic part; $\boldsymbol{\sigma}(\boldsymbol{X}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta}_1)d\mathbb{W}_t$ denotes the stochastic part of the system with

$$\boldsymbol{\sigma}(\cdot) = \text{diag}(\sigma_{free}C_{free}, \sigma_{con}C_{con}, \sigma_{sor}C_{sor}, \sigma_{seq}C_{seq}), \tag{3.13}$$

and $\boldsymbol{\theta}_1 = [k_{dc}, k_d, k_{de}, k_{bio}, k_{ab,free},\ k_{ab,sor}, \sigma_{free}, \sigma_{con}, \sigma_{sor}, \sigma_{seq}]$ denotes the vector of model parameters in the dynamic model (3.12). All other notation remains the same as previously defined in Section 3.2.1. It is worth mentioning that the ODE model (3.1-3.4) is the SDE model (3.8-3.11) with $\sigma_{free} = 0, \sigma_{con} = 0, \sigma_{sor} = 0, \sigma_{seq} = 0$.

3.2.3. Model for the observations

It is often the case that in wastewater treatment not all components of the system can be observed. Hence, the observational model is vital to statistically assess the disparity between the observed data and the model prediction. The observational model relates the dataset of external observations, $\boldsymbol{Y}$, to the internal state variables, $\boldsymbol{X}$. As described in section 2, three forms of EE2 measurements are available for this study. However, for the measurement with respect to EE2-parent concentration in solid phase, Y_3 it cannot be differentiated as to whether Y_3 represents:

$\mathcal{H}1$. the sorbed concentration, C_{sor},

$\mathcal{H}2$. the sequestered concentration, C_{seq},

$\mathcal{H}3$. the sum of a fraction of the sorbed parent compound together with the sequestered parent compound, $f_{sor}C_{sor} + C_{seq}$ with $0 \leq f_{sor} \leq 1$.

The aim is to elucidate which of these three hypotheses ($\mathcal{H}1$, $\mathcal{H}2$ and $\mathcal{H}3$) is most appropriate by fitting the data $\boldsymbol{Y}$ to the model (3.5) or (3.12). However, consideration shall only be given to model (3.12), since it is a more general model as discussed earlier in 3.2.2. The observational model can be written in the general framework as:

$$\boldsymbol{Y}_{n,q} = \boldsymbol{h}(\boldsymbol{X}_n, \boldsymbol{u}, t_n, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) + \boldsymbol{e}_{n,q}, \tag{3.14}$$

where t_n with $n = 1, 2, \cdots, N$ are the discrete time points; $(q = 1, 2, 3 = Q)$ is the number of repeated set; $\boldsymbol{Y}_{n,q} = [Y_1, Y_2, Y_3]^T_{n,q}$ are the experimentally measured concentrations at time point t_n for the qth experiment; $\boldsymbol{\theta}_2$ is the observational model parameter which will be defined in subsequent text and $\boldsymbol{h}(\cdot)$ denotes the observational function which we define here by equation (3.15) (see Text S1 in the supporting information for further details). Note,

$$\boldsymbol{h}(\cdot) = \begin{cases} \boldsymbol{h}_1(\cdot) & = [C_{free}, C_{con}, C_{sor}]^T_n \quad \text{for} \quad \mathcal{H}1 \\ \boldsymbol{h}_2(\cdot) & = [C_{free}, C_{con}, C_{seq}]^T_n \quad \text{for} \quad \mathcal{H}2 \\ \boldsymbol{h}_3(\cdot) & = [C_{free}, C_{con}, f_{sor}C_{sor} + C_{seq}]^T_n \quad \text{for} \quad \mathcal{H}3 \end{cases} \tag{3.15}$$

The unknown independent measurement errors at time point t_n for the qth experiment, $\boldsymbol{e}_{n,q}$ are assumed to belong to a multivariate normal distribution with mean zero and covariance matrix given by either

- a common measurement variance

$$\boldsymbol{\Sigma}^c_n = \text{diag}(\sigma_0^2, \sigma_0^2, \sigma_0^2) \tag{3.16}$$

- or different measurement variances,

$$\boldsymbol{\Sigma}^s_n = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2) \tag{3.17}$$

The covariance structure of the measurement noise given by equation (3.16) or (3.17) is called additive noise and it is the most fundamental noise structure considered for this study due to the nature of the experimental data (i.e. the observations at different times are assumed to have a common vari-

ance). Furthermore, the experimentally observed EE2 concentrations Y_1, Y_2 and Y_3 are assumed to be uncorrelated in all triplicate experimental data because of the experimental design used.

It is worth noting that hypothesis 3 ($\mathcal{H}3$) is identical to hypothesis 2 ($\mathcal{H}2$) when $f_{sor} = 0$. Hence, $\boldsymbol{\theta}_2 = [f_{sor}, \sigma_0^2]$ or $\boldsymbol{\theta}_2 = [f_{sor}, \sigma_1^2, \sigma_2^2, \sigma_3^2]$ is the additional parameter with respect to the observational model and so, the overall model parameter is $\boldsymbol{\theta} = \boldsymbol{\theta}_1 \cup \boldsymbol{\theta}_2$. Therefore, the observation model given by equation (3.14) becomes

$$\boldsymbol{Y}_{n,q} = \boldsymbol{h}_\varphi(\boldsymbol{X}_n, \boldsymbol{u}_n, t_n, \boldsymbol{\theta}) + \boldsymbol{e}_{n,q}, \quad \boldsymbol{e}_{n,q} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_n), \tag{3.18}$$

where $\boldsymbol{h}_\varphi$ with $\varphi = \{1, 2, 3\}$ is the observational function with respect to the three hypotheses $\mathcal{H}1$, $\mathcal{H}2$ and $\mathcal{H}3$.

4. Inference for dynamical models

4.1. Estimation model parameters

The aim is to fit $\boldsymbol{\theta} = [k_{dc}, k_d, k_{de}, k_{bio}, k_{ab,free}, k_{ab,sor}, f_{sor}, \sigma_{free}, \sigma_{con}, \sigma_{sor}, \sigma_{seq}, \sigma_0^2, \sigma_1^2, \sigma_2^2, \sigma_3^2]$ in the more general model which is the stochastic ODE as described in Sections 3.2.2 and 3.2.3. In this paper the method of parameter estimation of SDE proposed by Picchini et al. (2006) is applied to the problem under study, that is, equations (3.12) and (3.18). This is a method used to carry out statistical inference in discretely observed diffusion process with partially observed state variables. From (3.18), the simulated maximum likelihood function approximation of the conditional density function for each hypothesis, $\varphi = \{1, 2, 3\}$ is

$$L^R(\boldsymbol{\theta}|\boldsymbol{Y}) = \prod_{q=1}^{3} \left\{ \frac{1}{R} \sum_{r=1}^{R} \prod_{n=1}^{N-1} \phi_{n,q}(\boldsymbol{Y}_{n,q}|\boldsymbol{h}_\varphi(\boldsymbol{X}_n^{(r)}), \boldsymbol{\theta}) \right\}, \tag{4.1}$$

where, R denotes the number of Monte Carlo simulations, $\boldsymbol{X}_n^{(r)}$ denotes the rth simulated solution, $\boldsymbol{X}_n$ of the equation (3.12) using the Euler-Maruyama numerical scheme (Kloeden and Platen, 1992) at t_n for a given $\boldsymbol{\theta}$ and $\boldsymbol{X}_0$ at t_0 with the interval $[t_{n-1}, t_n]$ divided into M (large) subintervals of length $g = (t_n - t_{n-1})/M$. The multivariate normal density

function, $\phi_{n,q}(\boldsymbol{Y}_{n,q}|\boldsymbol{h}_\varphi(\boldsymbol{X}_n^{(r)}),\boldsymbol{\theta})$ with expectation $\boldsymbol{h}_\varphi(\boldsymbol{X}_n^{(r)})$ and covariance $\boldsymbol{\Sigma}_n(\boldsymbol{\theta})$ given by equations (3.16 and 3.17) is

$$\phi_{n,q}(\boldsymbol{Y}_{n,q}|\boldsymbol{h}_\varphi(\boldsymbol{X}_n^{(r)}),\boldsymbol{\theta}) = (2\pi)^{-3/2}(\det(\boldsymbol{\Sigma}_n(\boldsymbol{\theta})))^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{\epsilon}_{n,q}^T(\boldsymbol{\theta})\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\theta})\boldsymbol{\epsilon}_{n,q}(\boldsymbol{\theta})\right), \tag{4.2}$$

where, $\boldsymbol{\epsilon}_{n,q} = \boldsymbol{Y}_{n,q} - \boldsymbol{h}_\varphi(\boldsymbol{X}_n^{(r)})$ are the estimated errors at the time point t_n for the qth experiment. Therefore, the simulated log-likelihood function based on equation (4.1) ($\ell^R(\boldsymbol{\theta}|\boldsymbol{Y}) = \log L^R(\boldsymbol{\theta}|\boldsymbol{Y})$) is given by

$$\ell^R(\boldsymbol{\theta}|\boldsymbol{Y}) = \sum_{q=1}^{3}\left\{\log\left(\frac{1}{R}\sum_{r=1}^{R}\left[\prod_{n=1}^{N-1}\left((2\pi)^{-3/2}(\det(\boldsymbol{\Sigma}_n(\boldsymbol{\theta})))^{-1/2}\exp\left(-\frac{1}{2}\boldsymbol{\epsilon}_{n,q}^T(\boldsymbol{\theta})\boldsymbol{\Sigma}_n^{-1}(\boldsymbol{\theta})\boldsymbol{\epsilon}_{n,q}(\boldsymbol{\theta})\right)\right)\right]\right)\right\}. \tag{4.3}$$

The parameter estimates are obtained by maximising the simulated log-likelihood function (4.3) with respect to $\boldsymbol{\theta}$.

As described in Section 3.2, the three transformation pathways types we wish to compare for the drift term introduced in Section 3.2.2 within the stochastic modelling set up are ($\mathcal{M}0(\{\boldsymbol{\theta} : k_{ab,free} \neq 0, k_{ab,sor} \neq 0\})$, $\mathcal{M}1(\{\boldsymbol{\theta} : k_{ab,free} \neq 0, k_{ab,sor} = 0\})$ and $\mathcal{M}2(\{\boldsymbol{\theta} : k_{ab,free} = 0, k_{ab,sor} \neq 0\})$). For each pathway type, four levels of stochasticity (system noise) are considered as follows in order to capture any unknown variability that may exist:

- Full system noise, i.e., noise factor applied to all EE2 concentration variables ($\mathcal{S}1(\{\boldsymbol{\theta} : \sigma_{free} \neq 0, \sigma_{con} \neq 0, \sigma_{sor} \neq 0, \sigma_{seq} \neq 0\})$),
- Noise factor applied to all EE2 concentration variables except the sequestered EE2 concentration variable ($\mathcal{S}2(\{\boldsymbol{\theta} : \sigma_{free} \neq 0, \sigma_{con} \neq 0, \sigma_{sor} \neq 0, \sigma_{seq} = 0\})$),
- Noise factor applied to only conjugate EE2 concentration variable ($\mathcal{S}3(\{\boldsymbol{\theta} : \sigma_{free} = 0, \sigma_{con} \neq 0, \sigma_{sor} = 0, \sigma_{seq} = 0\})$),
- No system noise, i.e., fully deterministic model (3.5) ($\mathcal{S}4(\{\boldsymbol{\theta} : \sigma_{free} = 0, \sigma_{con} = 0, \sigma_{sor} = 0, \sigma_{seq} = 0\})$).

Note that for $\mathcal{S}4$, one does not need to simulate $\boldsymbol{X}_n$ using Monte Carlo simulations but rather one solves the ODE, for example via a Matlab ode

solver. To facilitate an efficient numerical search of the parameter space across a wide range of $\boldsymbol{\theta}$, the transformed parameter vector $log_{10}\boldsymbol{\theta}$ is estimated. Such estimates are robust since they are less prone to sampling error than when one directly works with untransformed parameters (Bland and Altman, 1996). The uncertainty in the parameter estimates obtained via approximate 95% confidence intervals are first computed on the *log*-scale before inverting the logarithm. The mathematical software used here is MATLAB, 2018a on an Intel(R) Core(TM) i5-8250U CPU @ 1.60Gh Hz @ 1.80Gh Hz Dell laptop. The pseudocode for implementing the simulated maximum likelihood parameter estimation is given in SDEMethod 1 in the supporting information.

The approximated maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ converges to the maximum likelihood estimator $\boldsymbol{\theta}$, where $\boldsymbol{\theta} = \boldsymbol{\theta}_1 \cup \boldsymbol{\theta}_2$ defined before in Section 3.2 as $M \to \infty$ and $R \to \infty$, with $R^{1/2}/M \to 0$ under certain mild regularity conditions (Brandt and Santa-Clara, 2002). To ensure that the value of the simulated maximum likelihood converges, we simulate equation (3.12) via the Euler-Maruyama scheme with an integration step-size of 0.0167 h respectively for $R = 1000, 2000, 3000, 4000$ and 5000 trajectories. The reported parameter estimates of $\boldsymbol{\theta}$ are based on $R = 3000$ since beyond this number of Monte-Carlo simulated trajectories, the value of the simulated likelihood function does not significantly change.

4.2. Model comparison

The likelihood ratio test (LRT) will be employed to compare two nested computational models in this paper (Madsen and Thyregod, 2010). Denoting the log-likelihood of the full and the reduced models by $\ell(\boldsymbol{\theta}_{\mathcal{H}_a})$ and $\ell(\boldsymbol{\theta}_{\mathcal{H}_0})$, the likelihood ratio test statistic is given by

$$\text{LRT} = 2(\ell(\boldsymbol{\theta}_{\mathcal{H}_a}) - \ell(\boldsymbol{\theta}_{\mathcal{H}_0})). \tag{4.4}$$

Under certain regularity conditions (Wilks, 1962), LRT has an asymptotic (χ^2) chi-square distribution with (df) degrees of freedom and df is the difference between the dimension of the vectors $\boldsymbol{\theta}_{\mathcal{H}_a}$ and $\boldsymbol{\theta}_{\mathcal{H}_0}$ for large sample sizes.

Alternatively, for the non-nested models the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) given by

$$\text{AIC} = 2(-\ell(\boldsymbol{\theta}) + p), \tag{4.5}$$

$$\text{BIC} = -2\ell(\boldsymbol{\theta}) + p \cdot \ln(w), \tag{4.6}$$

are employed, where p is the number of model parameters, $w = Q(N-1)k'$ denotes number of observations; k' denotes number of measured states and $\ell(\boldsymbol{\theta})$ is the model log-likelihood value. Often, the data available for use in studies involving wastewater treatment and receiving environments (e.g., rivers) are limited in size, so inferential studies are carried out on small sample sizes. For small sample sizes (w) such that $w/p < 40$, Burnham and Anderson (2003) suggested the use of the corrected AIC (AIC_C) given by

$$\text{AIC}_C = \text{AIC} + \frac{2p(p+1)}{w-p-1}. \tag{4.7}$$

In fisheries studies, the AIC_C has been shown to be accurate (Shono, 2000; Katsanevakis, 2006). Other related studies have also shown that the use of these information criteria is useful in comparing models (Shukor, 2014; Razzaghi et al., 2016; Mazerolle, 2006; Burnham et al., 2011). When comparing models, the one with the lower AIC, AIC_C and BIC value is preferred.

5. Results and discussion

5.1. Experimental data

Figure 3 shows the triplicate experimental profiles of EE2-parent (Y_1), -conjugated (Y_2) and -solid (Y_3) and these profiles are similar. This gives an indication of low variability between the triplicates in the experimental set-up, and implies consistency in the experimental method since variations in the concentrations of the observations will be due to field conditions rather than experimental design.

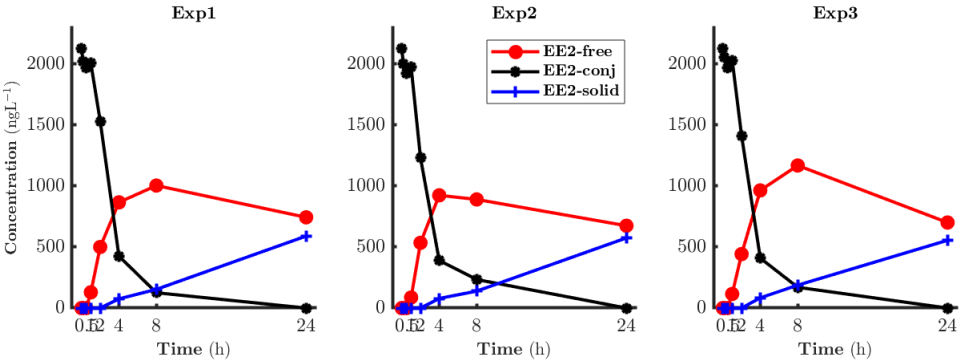


Fig. 3. Profiles of liquid EE2-parent, EE2-conjugated and solid EE2-parent, respectively in triplicate experimental set-up: Experiment 1 (Expt1) (**left**), Experiment 2 (Expt2) (**middle**) and Experiment 3 (Expt3) (**right**) plotted at times $0, 0.167, 0.5, 1, 2, 4, 8, 10$ and 24 h. Experimental data points, denoted by red circle, black asterisk and blue plus, respectively shows liquid EE2-parent (Y_1), liquid retransformable EE2 (Y_2) and solid EE2-parent (Y_3)(adapted from Gomes et al. (2009)).

As can be seen in Figure 3 (all panels), over the first 4 hours, there is a rapid decrease in the EE2-conjugate concentration, after which the rate of reduction slows considerably. The EE2-conjugate is converted to liquid EE2-parent (free), which increases rapidly from 30 mins to 8 hours and then slowly declines. The reason for zero concentrations being observed at $t = 10$ mins and $t = 30$ mins is that the measured concentrations may be below the analytical limit of detection, and so have been recorded as zero. Whilst both of the above concentrations show transient behaviour, the EE2-parent concentration in the solid phase (adsorbed or sequestered) does not, exhibiting only a steady slow growth over the 24 hours study period. Whilst the concentration in the solid phase was recorded as zero for times up to 2 hours, it may have been larger, but below the limit of detection.

The question this study aims to address is whether one can associate the steady slow growth trend for solid EE2-parent to adsorption or to absorption or to both. Since there is an adsorption - desorption process between liquid EE2-parent and sorbed EE2-parent, one may expect the profile of solid EE2-parent to be similar to that of liquid EE2-parent but that is not the case here.

5.2. Inference for dynamical model

5.2.1. Model parameter estimation

The simulated log-likelihood estimates computed from equation (4.3) under hypotheses $\mathcal{H}1, \mathcal{H}2$ and $\mathcal{H}3$ are presented in Table 2 for the cases of common observational variance (3.16) and separate variances (3.17); various diffusion terms ($\mathcal{S}1, \mathcal{S}2, \mathcal{S}3$ and $\mathcal{S}4$) and model structures ($\mathcal{M}0, \mathcal{M}1$ and $\mathcal{M}2$), see Sections Sections 3.2 and 4.1. The notion of common and separate observational noise is tested for each hypothesis ($\mathcal{H}1$, $\mathcal{H}2$ and $\mathcal{H}3$) with respect to the three system noises as well as the three model structures using the likelihood ratio test, AIC, AIC_C and BIC as given in Section 4.2.

In this study, the model parameters to be estimated can be grouped into three parts, those associated with the drift term, $\boldsymbol{f}(\cdot)dt$ in (3.12), diffusion term, $\boldsymbol{\sigma}(\cdot)d\mathbb{W}$ in (3.12), and the observational noise, $\boldsymbol{e}$ in (3.18) as described in Section 3.2.3. As a result, hypotheses $\mathcal{H}1$ and $\mathcal{H}2$ have from $p = 6$ parameters for model $\mathcal{M}2$ under $\mathcal{S}4$ to $p = 11$ parameters for model $\mathcal{M}0$ under $\mathcal{S}1$ to be estimated, whilst under hypothesis $\mathcal{H}3$ the number of parameters ranges from $p = 7$ for model $\mathcal{M}2$ under $\mathcal{S}4$ to $p = 12$ parameters for model $\mathcal{M}0$ under $\mathcal{S}1$, when a common observational noise (3.16) is considered. When separate observational noise (3.17) is considered, the total number of parameters ranges from $p = 8$ for $\mathcal{M}2$ under $\mathcal{S}4$ to $p = 13$ for $\mathcal{M}0$ under $\mathcal{S}1$ for hypotheses $\mathcal{H}1$ and $\mathcal{H}2$ whilst hypothesis $\mathcal{H}3$ have parameters ranges from $p = 9$ for $\mathcal{M}2$ under $\mathcal{S}4$ to $p = 14$ for $\mathcal{M}0$ under $\mathcal{S}1$. For this reason, the number of degrees of freedom, df, between any two nested models ranges from to 1 to 5. Thus, the critical values for at χ^2_{df} with df degrees of freedom at significance level of $\alpha = 0.05$, $\chi^2_{df}(0.05)$ are $3.8415, 5.9915, 7.8147, 9.4877$ and 11.0704 respectively for $df = 1, 2, 3, 4$ and 5.

For each hypothesis, separate additive observational noise is tested against common additive noise, and it is found that a significance difference exists (i.e., separate additive observational noise is better than a common observational additive noise) across all models under hypotheses $\mathcal{H}1, \mathcal{H}2$ and $\mathcal{H}3$. This conclusion is based on the estimates LRT, AIC, AIC_C and BIC which give similar results under hypotheses $\mathcal{H}1, \mathcal{H}2$ and $\mathcal{H}3$. Altogether this gives an indication that a separate additive noise better describes the experimen-

Table 2

Results of model parameter estimation: negative log-likelihood, AIC, AIC_C and BIC for each of the three hypotheses ($\mathcal{H}1\,(Y_3 = C_{sor}), \mathcal{H}2\,(Y_3 = C_{seq})$ and $\mathcal{H}3\,(Y_3 = C_{sor} + f_{sor}C_{seq}, 0 \leq f_{sor} \leq 1)$) with respect to three model structures (MS) of $\boldsymbol{f}(\cdot)((\mathcal{M}0(k_{ab,free} \neq 0, k_{ab,sor} \neq 0), \mathcal{M}1(k_{ab,free} \neq 0, k_{ab,sor} = 0)$ and $\mathcal{M}2(k_{ab,free} = 0, k_{ab,sor} \neq 0)))$ as well as the system noise (SN) ($\mathcal{S}1(\sigma_{free} \neq 0, \sigma_{con} \neq 0, \sigma_{sor} \neq 0, \sigma_{seq} \neq 0), \mathcal{S}2(\sigma_{free} \neq 0, \sigma_{con} \neq 0, \sigma_{sor} \neq 0, \sigma_{seq} = 0), \mathcal{S}3(\sigma_{free} = 0, \sigma_{con} \neq 0, \sigma_{sor} = 0, \sigma_{seq} = 0, \sigma_{sor} \neq 0, \sigma_{seq} = 0), \mathcal{S}4(\sigma_{free} = 0, \sigma_{con} = 0, \sigma_{sor} = 0, \sigma_{seq} = 0)$) in relation to sequestration formation and p is the number of model parameters. LRT denotes likelihood ratio estimates between models with common and separate additive observational noise with corresponding p-values less than 10^{-6}. Red highlighted values denotes overall best fit models.

			Common ($\sigma_1 = \sigma_2 = \sigma_3 = \sigma_0$)					Separate ($\sigma_1 = \sigma_2 = \sigma_3$)					
Hypothesis	SN	MS	p	$\ell(\boldsymbol{\theta})$	AIC	AIC_C	BIC	p	$\ell(\boldsymbol{\theta})$	AIC	AIC_C	BIC	LRT
$\mathcal{H}1$	$\mathcal{S}1$	$\mathcal{M}0$	11	-398.96	819.92	825.09	843.49	13	-361.32	748.64	756.07	776.50	75.28
		$\mathcal{M}1$	10	-398.94	817.87	822.11	839.31	12	-361.03	746.06	752.30	771.78	75.82
		$\mathcal{M}2$	10	-398.96	817.92	822.15	839.35	12	-361.46	746.92	753.16	772.64	75.10
	$\mathcal{S}2$	$\mathcal{M}0$	10	-398.96	817.92	822.15	839.35	12	-361.96	747.93	754.17	773.64	74.00
		$\mathcal{M}1$	9	-398.94	815.87	819.27	835.16	11	-361.05	744.09	749.27	767.67	75.78
		$\mathcal{M}2$	9	-398.96	815.92	819.31	835.20	11	-361.09	744.18	749.36	767.76	75.74
	$\mathcal{S}3$	$\mathcal{M}0$	8	-398.91	813.82	816.49	830.96	10	-360.90	741.81	746.64	763.24	76.02
		$\mathcal{M}1$	7	-398.65	811.30	813.34	826.31	9	-360.63	739.26	742.66	758.55	76.04
		$\mathcal{M}2$	7	-398.67	811.34	813.38	826.35	9	-360.73	739.46	742.86	758.75	75.88
	$\mathcal{S}4$	$\mathcal{M}0$	7	-413.36	840.72	842.76	855.72	9	-368.62	755.24	758.64	832.04	89.48
		$\mathcal{M}1$	6	-413.34	838.69	840.19	851.54	8	-368.56	753.12	755.79	828.77	89.56
		$\mathcal{M}2$	6	-413.36	838.72	840.22	851.58	8	-368.58	753.16	755.82	827.98	89.56
$\mathcal{H}2$	$\mathcal{S}1$	$\mathcal{M}0$	11	-373.94	769.89	775.06	793.46	13	-342.66	711.32	718.75	739.18	62.56
		$\mathcal{M}1$	10	-374.14	768.27	772.50	789.70	12	-342.99	709.98	716.22	735.70	62.3
		$\mathcal{M}2$	10	-373.92	767.85	772.08	789.28	12	-342.66	709.32	715.56	735.04	62.52
	$\mathcal{S}2$	$\mathcal{M}0$	10	-373.90	767.80	772.03	789.23	12	-342.73	709.47	715.71	735.18	62.34
		$\mathcal{M}1$	9	-373.90	765.80	769.19	785.09	11	-342.83	707.66	712.84	731.23	62.14
		$\mathcal{M}2$	9	-373.90	765.81	769.20	785.09	11	-342.69	707.38	712.56	730.95	62.42
	$\mathcal{S}3$	$\mathcal{M}0$	8	-374.77	765.55	768.21	782.69	**10**	**-342.94**	**705.88**	**710.11**	**727.31**	63.66
		$\mathcal{M}1$	7	-374.73	763.46	765.49	778.46	**9**	**-342.91**	**703.82**	**707.22**	**723.11**	63.64
		$\mathcal{M}2$	7	-374.63	763.25	765.29	778.25	**9**	**-342.89**	**703.77**	**707.17**	**723.06**	63.48
	$\mathcal{S}4$	$\mathcal{M}0$	7	-399.48	812.96	815.00	827.96	9	-358.14	734.29	737.69	753.58	82.68
		$\mathcal{M}1$	6	-399.77	811.55	813.05	824.41	8	-359.87	735.74	738.40	752.88	79.80
		$\mathcal{M}2$	6	-399.57	811.13	812.63	823.99	8	-357.93	731.87	734.53	749.01	83.28
$\mathcal{H}3$	$\mathcal{S}1$	$\mathcal{M}0$	12	-374.09	772.12	778.41	797.89	14	-359.43	746.86	755.61	776.87	29.32
		$\mathcal{M}1$	11	-373.93	769.85	775.03	793.43	13	-342.71	711.42	718.85	739.28	62.44
		$\mathcal{M}2$	11	-373.87	769.75	774.92	793.32	13	-342.99	711.99	719.42	739.85	61.76
	$\mathcal{S}2$	$\mathcal{M}0$	11	-373.87	769.74	774.91	793.31	13	-342.76	711.51	718.94	739.38	62.22
		$\mathcal{M}1$	10	-373.85	767.70	771.93	789.13	12	-343.13	710.25	716.49	735.97	61.44
		$\mathcal{M}2$	10	-373.85	767.71	771.94	789.14	12	-342.78	709.56	715.80	735.28	62.14
	$\mathcal{S}3$	$\mathcal{M}0$	9	-374.59	767.17	770.57	786.46	11	-343.00	708.01	713.18	731.58	63.18
		$\mathcal{M}1$	8	-374.61	765.23	767.89	782.37	10	-343.23	706.46	710.69	727.89	62.76
		$\mathcal{M}2$	8	-374.76	765.52	768.19	782.67	10	-343.02	706.03	710.26	727.46	63.48
	$\mathcal{S}4$	$\mathcal{M}0$	8	-399.45	814.89	817.56	832.04	10	-358.90	737.79	742.03	759.23	81.1
		$\mathcal{M}1$	7	-399.88	813.77	815.81	828.77	9	-357.98	733.96	737.36	753.25	83.80
		$\mathcal{M}2$	7	-399.49	812.97	815.01	827.98	9	-357.93	733.86	737.25	753.14	83.12

tal data. This might imply that the assumption where the measurement errors have a common variance (3.16) may not hold, as indicated by the results for lab-scale data.

The analysis of the models is further extended to examine the hypotheses $\mathcal{H}1, \mathcal{H}2$ and $\mathcal{H}3$ with respect to separate additive observational noise. Results from Table 2 clearly show that between hypotheses $\mathcal{H}1$ and $\mathcal{H}2$, models under hypothesis $\mathcal{H}2$ are significantly better based on the values of $\ell(\boldsymbol{\theta})$, AIC, AIC_C and BIC. On the other hand, applying the likelihood ratio test (LRT) in equation (4.4) to hypotheses $\mathcal{H}2$ and $\mathcal{H}3$ provides no evidence against the simpler models of hypothesis $\mathcal{H}2$. These results are supported by the corresponding values of AIC, AIC_C and BIC given by equations (4.5) - (4.6). Since C_{sor} and C_{seq} are usually indistinguishable, one might expect hypothesis $\mathcal{H}3$ to hold, however results from the fit suggest that hypothesis $\mathcal{H}2$ best describes the data. This outcome confirms previous findings on sequestration and in addition, quantified sequestered EE2 from the modelling perspective (Barret et al., 2011; Banihashemi and Droste, 2014).

The results of the simulated maximum likelihood estimation of the selected models of hypothesis $\mathcal{H}2$ and $\mathcal{S}3$ are shown in Table 3 and the dynamical simulated results with these parameters are depicted in Figure 4. The estimated value of the diffusion parameter or the system noise associated with the conjugate EE2 state variable (C_{con}), σ_{con} is similar across all the three model structures considered ($\mathcal{M}0, \mathcal{M}1$ and $\mathcal{M}2$). Statistically testing these mixed deterministic-stochastic models under $\mathcal{S}3$ with purely deterministic models under $\mathcal{S}4$, suggests that there are strong evidence for rejecting the purely deterministic models since the p-values corresponding to these tests are less than 10^{-7}. Furthermore referring to Table 3, it can be observed that the difference between their estimated values of AIC, AIC_C and BIC are small, thereby indicating that all these models adequately describe the experimental data. This is also evident in their simulated trajectories illustrated in Figure 4. Interestingly the model fits of $\mathcal{M}0$, $\mathcal{M}1$ and $\mathcal{M}2$ under hypothesis $\mathcal{H}2$ appear similar from the graphical point of view (see Figure 4).

Table 3
Simulated maximum likelihood (SML) parameter estimate with 95% confidence intervals in parenthesis for hypothesis $\mathcal{H}2$ under separate observational noise with respect to model structures (transformation pathways types) $\boldsymbol{f}(\cdot)$ $\mathcal{M}0(k_{ab,free} \neq 0, k_{ab,sor} \neq 0)$, $\mathcal{M}1(k_{ab,free} \neq 0, k_{ab,sor} = 0)$ and $\mathcal{M}0(k_{ab,free} = 0, k_{ab,sor} \neq 0)$

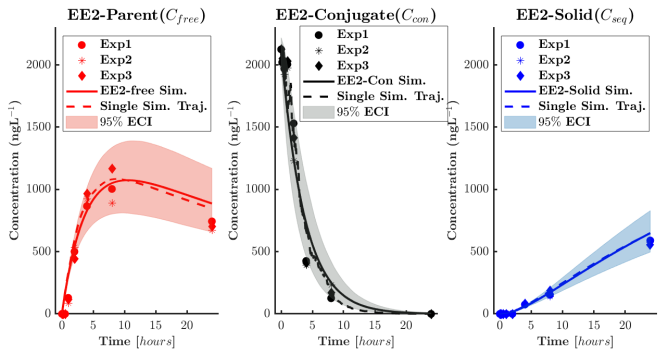
System noise 3 ($\mathcal{S}3(\sigma_{free} = 0, \sigma_{con} \neq 0, \sigma_{sor} = 0, \sigma_{seq} = 0)$)				
	$\mathcal{M}0$	$\mathcal{M}1$	$\mathcal{M}2$	
Param.	$\hat{\boldsymbol{\theta}}$ (95%CI)	$\hat{\boldsymbol{\theta}}_R$ (95%CI)	$\hat{\boldsymbol{\theta}}_R$ (95%CI)	Units
k_{dc}	**0.423**	**0.425**	**0.424**	h^{-1}
	$(0.380, 0.471)$	$(0.384, 0.470)$	$(0.383, 0.470)$	
k_{sor}	**11.242**	**11.460**	**11.430**	LgX_{SS}^{-1}
	$(8.781, 14.392)$	$(9.352, 14.043)$	$(9.352, 13.970)$	
k_{de}	**74.828**	**73.963**	**74.038**	h^{-1}
	$(64.382, 86.968)$	$(65.315, 83.756)$	$(65.634, 83.517)$	
k_{bio}	$\mathbf{5.020 \times 10^{-4}}$	$\mathbf{5.123 \times 10^{-4}}$	$\mathbf{5.077 \times 10^{-4}}$	$\mathrm{Lg}^{-1}\mathrm{h}^{-1}$
	$(1.834 \times 10^{-6}, 0.137)$	$(1.161 \times 10^{-5}, 0.023)$	$(9.481 \times 10^{-6}, 0.027)$	
$k_{ab,free}$	**0.025**	**0.030**	–	h^{-1}
	$(0.016, 0.040)$	$(0.028, 0.033)$	–	
$k_{ab,sor}$	**0.007**	–	**0.048**	h^{-1}
	$(5.657 \times 10^{-4}, 0.096)$	–	$(0.038, 0.062)$	
σ_{con}	**0.127**	**0.128**	**0.129**	–
	$(0.098, 0.164)$	$(0.101, 0.164)$	$(0.101, 0.164)$	
σ_1	**126.430**	**127.350**	**128.290**	ng
	$(93.091, 171.710)$	$(93.671, 173.140)$	$(93.676, 173.210)$	
σ_2	**96.845**	**96.977**	**97.056**	ng
	$(70.967, 132.160)$	$(71.175, 132.130)$	$(71.207, 132.290)$	
σ_3	**14.294**	**14.169**	**14.28**	ng
	$(10.520, 19.421)$	$(10.464, 19.187)$	$(10.519, 19.387)$	
$\ell(\hat{\boldsymbol{\theta}})$	-342.940	-342.910	-342.890	–
AIC	705.880	703.820	703.770	–
AIC_C	710.110	707.220	707.170	–
BIC	727.310	723.110	723.060	–

Figure 4 shows the experimental observed data undertaken in triplicate at each time point and the empirical mean of $R = 5000$ simulated trajectories of the model outputs, together with their associated empirical 95% pointwise confidence limits. It includes all the three hypothesis with respect to the models (transformation pathways types) $\mathcal{M}0$, $\mathcal{M}1$ and $\mathcal{M}2$ under hypothesis $\mathcal{H}2$ and $\mathcal{S}3$, where the experimental measured EE2-parent concentration in the solid phase (Y_3) is solely the sequestered EE2-parent ($X_3 = C_{seq}$). In

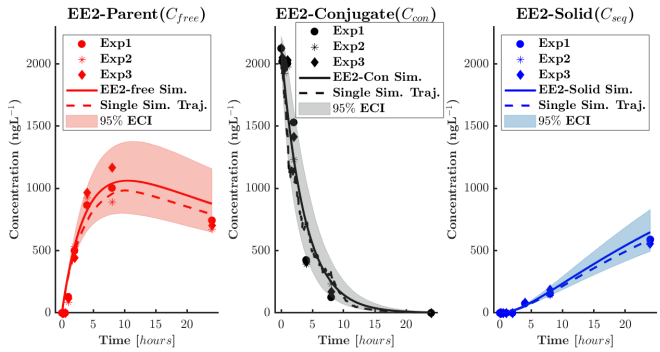
all the cases, the deconjugation of the EE2-conjugate is associated with an increase in the liquid EE2-parent concentration over the first 8 hours, after which the liquid EE2-parent concentration begins to decrease as a result of the combination of adsorption, sequestration and biodegradation processes while EE2-conjugate decreases to zero and would be expected (Johnson and Williams, 2004). The liquid EE2-parent concentration sequestered is associated with an increase in the EE2-parent concentration in the solid phase observed in Figures 3 and 4.

The EE2 SDE models $\mathcal{M}0, \mathcal{M}1$ and $\mathcal{M}2$ under hypothesis $\mathcal{H}2$ with respect to system noise type $\mathcal{S}3$ capture the experimental data well over the entire 24 hours and, in addition, the empirical 95% confidence limits contain almost all the experimental data. From Figure 4, the uncertainty represented by the 95% empirical confidence limits especially for EE2-conjugate shows that the multiplicative noise in the SDE model accounts for influential factors that the deterministic ODE models cannot capture (not shown here). This is evident in the parameter estimate of σ_{con} given in Table 3. Therefore, the SDE model could account for certain inherent influential factors that are neglected by the ODE model (results under $\mathcal{S}4$). However, it is worth noting that only one of the SDE models is subject to noise (dC_{con}) and the other models are deterministic in the chosen model.

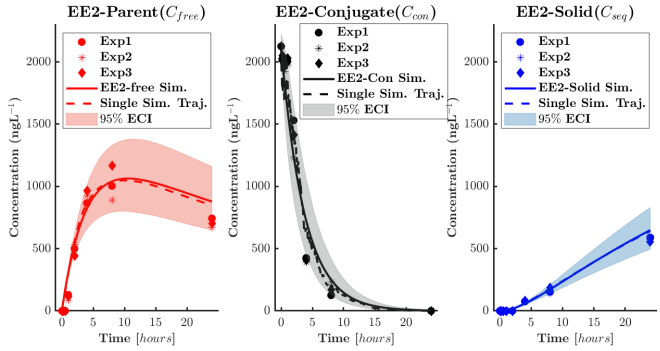
Figure S1 in the supporting information depicts histograms of the stochastic trajectories of $R = 5000$ realisations of the EE2 state variable concentrations (liquid EE2-parent, EE2-conjugate, and solid EE2-parent) for model $\mathcal{M}0$ under hypothesis $\mathcal{H}2$. The distribution of the EE2 state variable concentrations simulated at each time appears symmetrical, except for the EE2 state variable (conjugate) which is skewed at the last time point (row 7 of Figure S1). The skewness of EE2-conjugated at the last time point is not surprising because the measured EE2-conjugate concentration is zero at that time point and so the frequency of zero predicted concentration estimate by the SDE model is high. Overall, the histogram is approximately normal and so the model is acceptable. Similar distributions of each of the state variables in models $\mathcal{M}1$ and $\mathcal{M}2$ are respectively reported in Figures S2 and S3 in the supporting information.



(a) $\mathcal{M}0$



(b) $\mathcal{M}1$



(c) $\mathcal{M}2$

Fig. 4. Dynamics of EE2 SDE model showing experimental data, empirical mean, empirical 95% confidence limits with $R = 5000$ and a single trajectory under hypotheses $\mathcal{H}2$ for the models $\mathcal{M}0$ (***top***), $\mathcal{M}1$ (***middle***) and $\mathcal{M}2$ (***bottom***) with respect to $\mathcal{S}3$. Each panel has three sub panels of each state variable dynamics, namely: EE2- Parent (***left***), EE2-Conjugate (***middle***) and EE2-Solid (***right***) plotted at times $0, 0.167, 0.5, 1, 2, 8$ and 24 h.

The similarities of the distributions of each of the state variables in the models $\mathcal{M}0$, $\mathcal{M}1$ and $\mathcal{M}2$ plotted at the time points under hypothesis $\mathcal{H}2$ shows that there is no difference between $\mathcal{M}0$, $\mathcal{M}1$ and $\mathcal{M}2$ as indicated by the likelihood ratio test and information criteria, and so $\mathcal{M}0, \mathcal{M}1$ and $\mathcal{M}2$ are taken as plausible pathways. This suggests that more data are needed to distinguish these pathways. What can be said however is that, there is evidence of sequestration (Barret et al., 2011) and that the measured EE2-parent concentration in the solid phase, Y_3 corresponds to the EE2-sequestered, C_{seq} (which is hypothesis $\mathcal{H}2$) for the experimental dataset (Gomes et al., 2009) used in this study. This is surprising, considering that in the lab and field, these two types of sorption to the solid phase (C_{sor} and C_{seq}) are rarely defined separately (Joss et al., 2006). However, from the modelling point of view the solid (sorbed) parent compound (EE2), C_{sor} equilibrates with the liquid parent compound C_{free} so fast that the solid phase which is measured in the lab rather is the sequestered EE2 concentration in the solid phase (i.e., $Y_3 = C_{seq}$) as illustrated in Figure 4. Thus, C_{seq} is an important novel component in the model.

5.3. Perspective of ODE and SDE models

In general, the rich methodologies of both deterministic and stochastic modelling frameworks have been utilised to gain useful insights into the proposed models (Figure 2) and to compare both approaches. The stochastic modelling results turn out to be significantly better (see Table 2) and further allows the distinction between the model system noise which is captured by the parameter σ_{con} (accounting for inherent randomness of deconjugation of metabolites) and observational randomness accounted for in this study by separate additive noise.

Furthermore, the results of the parameter estimates obtained in this study show that not accounting for certain dynamics may lead to over- or under-estimation of model parameters which may contribute to variation in observed concentrations of emerging pollutants, as pointed out by authors (Liu et al., 2015; Barret et al., 2011). An example is the deconjugation kinetic rate coefficient, k_{dc} obtained by Gomes et al. (2009) which is 0.32

h^{-1} and is less than the estimated value reported in Table 3 under hypothesis $\mathcal{H}2$ where we find $k_{dc} = 0.42\ h^{-1}$. Differences in kinetic rate estimates in this paper to others are not surprising since the model accounts for sequestration dynamics, which to the best of the authors' knowledge has to date been neglected in the literature.

5.4. Significance

This present study investigates the mechanism of sorption of EE2 to the solid phase in an activated sludge process. In addition, the study examines the nature of the measured EE2 concentration in the solid phase as to whether it denotes the sorbed EE2 or sequestered EE2 or a combination of both from the modelling perspective. This study demonstrates and confirms sequestration, a phenomena known in literature yet rarely considered or distinguished beyond a pollutant associated with the solid phase (be that MLSS, sediment, sludge) (Barret et al., 2011; Yi et al., 2006; Yi and Harper, 2007). Results from the modelling strongly suggests that the experimentally measured EE2 concentration in the solid phase is the sequestered EE2, which is explained by hypothesis $\mathcal{H}2$. In particular, for any BACs having high affinity for a solid phase like EE2 there may be similar behaviour because of similarities in their physiochemical properties.

This study shows the usefulness of mathematical modelling in wastewater treatment as the paper evaluates sequestration of EE2, and this will inform future efforts to develop and as well as apply models better able to fit and/or predict BAC behaviour and fate in the environment. The challenge, however, is that most data available from the literature does not capture the changes in the concentrations over time, where all of the three BAC states (e.g., liquid EE2-parent, solid EE2-parent and EE2-conjugate) are determined simultaneously. In the future, experiments of such nature may be encouraged to enable modelling as a route to better understand wastewater treatment process performance. Specifically, sequestered BAC concentrations will need to be measured as recommended in the literature (Barret et al., 2011).

6. Conclusions

Mathematical modelling of emerging pollutants in wastewater treatment offers a way of investigating phenomena and establishing which mechanisms affect the fate and behaviour of BACs in WWTPs. This study, has performed a proof-of-concept investigation of EE2 fate in an activated sludge pilot plant. In addition, the study examined the nature of the measured EE2 concentration in the solid phase via deterministic and stochastic modelling frameworks to gain insight into the fate and behaviour of emerging pollutants such as bioactive compounds, with a specific case study of EE2. Proposed models have been tested using the likelihood ratio test (LRT), Akaike Information Criterion (AIC), corrected AIC (AIC_C) and Bayesian Information Criterion (BIC) to examine the nature of the experimental measured EE2 concentration in the solid phase from a lab-scale batch experiment.

From the computational models developed in this paper, the proposed pathway of EE2 in wastewater treatment describes the experimental dataset well, thus confirming the existence of sequestration for the EE2 dataset. In addition, the results suggest that the experimental measured EE2 concentration in the solid phase is more likely to be the sequestered parent compound (C_{seq}, hypothesis $\mathcal{H}2$) rather than either the sorbed parent compound (C_{sor}, hypothesis $\mathcal{H}1$) or the sum of a fraction of the sorbed parent compound together with the sequestered parent compound ($f_{sor}C_{sor} + C_{seq}$, hypothesis $\mathcal{H}3$). Furthermore, the results based on information criteria in Section 4.2 show that purely deterministic models do not adequately fit the data and stochastic models of the process are required. However, not all components need to be treated stochastically, and this work indicates that the best fit was obtained when just one of the four variables (C_{sor}) was stochastic. This work shows that making the distinction between adsorbed and absorbed (sequestered) pollutants is beneficial. It can be hypothesized that fate studies involving pollutants that prefer the solid phase should evaluate sequestration phenomena in order to provide a more accurate description of the process. The environmental consequence of this conclusion is that the uncertainty in EE2 or other BACs removal efficiency will be reduced, since wastewa-

ter treatment can be further optimized with this insight and this will aid stakeholders in decision making.

Acknowledgements

The authors (EA, RLG, ILD and JADW) are thankful for funding provided by the Leverhulme Trust Doctoral Scholarship (DA214-024), Modelling and Analytics for a Sustainable Society (MASS), University of Nottingham. EA is thankful for funding from Methodist University College Ghana (MUCG). EA gratefully acknowledges Prof. Krist V. Genaey, Dr. Flores Alsina, Dr. Gürkan Sin and Dr. Pedram Ramin for their support and engagement during and after his research visit and training at Denmark Technical University (DTU).

Appendix A. Supplementary data

Supplementary information related to this paper is attached as EA-manuscriptsuppliRV.pdf.

References

Banihashemi, B., Droste, R.L., 2014. Sorption–desorption and biosorption of bisphenol a, triclosan, and 17α-ethinylestradiol to sewage sludge. Science of the Total Environment 487, 813–821.

Barbosa, M.O., Moreira, N.F., Ribeiro, A.R., Pereira, M.F., Silva, A.M., 2016. Occurrence and removal of organic micropollutants: an overview of the watch list of eu decision 2015/495. Water Research 94, 257–279.

Barret, M., Carrère, H., Patau, M., Patureau, D., 2011. Kinetics and reversibility of micropollutant sorption in sludge. Journal of Environmental Monitoring 13, 2770–2774.

Bland, J.M., Altman, D.G., 1996. Statistics notes: measurement error proportional to the mean. British Medical Journal 313, 106.

Brandt, M.W., Santa-Clara, P., 2002. Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. Journal of Financial Economics 63, 161–210.

Burnham, K.P., Anderson, D.R., 2003. Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media.

Burnham, K.P., Anderson, D.R., Huyvaert, K.P., 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behavioral Ecology and Sociobiology 65, 23–35.

Carvalho, R.N., Ceriani, L., Ippolito, A., Lettieri, T., 2015. Development of the first watch list under the environmental quality standards directive. Sustainability EEJRCIfEa. Publications Office of the European Union, Luxembourg 166.

Coleman, M.D., 2010. Human drug metabolism. An Introduction. Second Edition. A John Wiley & Sons, Ltd., Publication .

Gernaey, K.V., van Loosdrecht, M.C., Henze, M., Lind, M., Jørgensen, S.B., 2004. Activated sludge wastewater treatment plant modelling and simulation: state of the art. Environmental Modelling & Software 19, 763–783.

Gomes, R.L., Scrimshaw, M.D., Lester, J.N., 2009. Fate of conjugated natural and synthetic steroid estrogens in crude sewage and activated sludge batch studies. Environmental Science & Technology 43, 3612–3618.

Johnson, A.C., Williams, R.J., 2004. A model to estimate influent and effluent concentrations of estradiol, estrone, and ethinylestradiol at sewage treatment works. Environmental Science & Technology 38, 3649–3658.

Joss, A., Zabczynski, S., Göbel, A., Hoffmann, B., Löffler, D., McArdell, C.S., Ternes, T.A., Thomsen, A., Siegrist, H., 2006. Biological degradation of pharmaceuticals in municipal wastewater treatment: proposing a classification scheme. Water Research 40, 1686–1696.

Katsanevakis, S., 2006. Modelling fish growth: model selection, multi-model inference and model selection uncertainty. Fisheries Research 81, 229–235.

Kloeden, P.E., Platen, E., 1992. Stochastic differential equations, in: Numerical Solution of Stochastic Differential Equations. Springer, Berlin, Heidelberg, pp. 103–160.

Kumar, V., Johnson, A.C., Nakada, N., Yamashita, N., Tanaka, H., 2012. De-conjugation behavior of conjugated estrogens in the raw sewage, activated sludge and river water. Journal of Hazardous Materials 227, 49–54.

Laurenson, J.P., Bloom, R.A., Page, S., Sadrieh, N., 2014. Ethinyl estradiol and other human pharmaceutical estrogens in the aquatic environment: a review of recent risk assessment data. The AAPS journal 16, 299–310.

Liu, Z.h., Lu, G.n., Yin, H., Dang, Z., Rittmann, B., 2015. Removal of natural estrogens and their conjugates in municipal wastewater treatment plants: a critical review. Environmental Science & Technology 49, 5288–5300.

Madsen, H., Thyregod, P., 2010. Introduction to general and generalized linear models. CRC Press.

Margot, J., Rossi, L., Barry, D.A., Holliger, C., 2015. A review of the fate of micropollutants in wastewater treatment plants. Wiley Interdisciplinary Reviews: Water 2, 457–487.

Mazerolle, M.J., 2006. Improving data analysis in herpetology: using Akaike's information criterion (AIC) to assess the strength of biological hypotheses. Amphibia-Reptilia 27, 169–180.

OECD, P.B.T., 2001. OECD guideline for the testing of chemicals. The Hershberger 601, 858.

Øksendal, B., 2003. Stochastic differential equations. Springer, Berlin, Heidelberg.

Picchini, U., Ditlevsen, S., De Gaetano, A., 2006. Modeling the euglycemic hyperinsulinemic clamp by stochastic differential equations. Journal of Mathematical Biology 53, 771–796.

Plósz, B., Lehnberg, K., Dott, W., Thomas, K., 2010. An activated sludge model for xenobiotic organic micro-pollutants (ASM-X): factors influencing the solid-liquid partitioning in wastewater treatment, in: 2nd IWA/WEF Wastewater Treatment Modelling Seminar, WWTmod 2010, Quebec, Canada March 28-30, 2010, pp. 297–305.

Plósz, B.G., Langford, K.H., Thomas, K.V., 2012a. An activated sludge modeling framework for xenobiotic trace chemicals (ASM-X): Assessment of diclofenac and carbamazepine. Biotechnology and Bioengineering 109, 2757–2769.

Plósz, B.G., Nopens, I., Rieger, L., Griborio, A., De Clercq, J., Vanrolleghem, P.A., Daigger, G.T., Takács, I., Wicks, J., Ekama, G.A., 2012b. A critical review of clarifier modeling: State-of-the-art and engineering practices, in: Proc. 3rd IWA/WEF Wastewater Treatment Modeling Seminar (WWTmod2012)]. Mont-Sainte-Anne, Quebec, pp. 27–30.

Polesel, F., Lehnberg, K., Dott, W., Trapp, S., Thomas, K.V., Plósz, B.G., 2015. Factors influencing sorption of ciprofloxacin onto activated sludge: Experimental assessment and modelling implications. Chemosphere 119, 105–111.

Razzaghi, S., Khodaverdiloo, H., Ghorbani Dashtaki, S., 2016. Effects of long-term wastewater irrigation on soil physical properties and performance of selected infiltration models in a semi-arid region. Hydrological Sciences Journal 61, 1778–1790.

Salgado, R., Marques, R., Noronha, J., Carvalho, G., Oehmen, A., Reis, M., 2012. Assessing the removal of pharmaceuticals and personal care products in a full-scale activated sludge plant. Environmental Science and Pollution Research 19, 1818–1827.

Sathyamoorthy, S., Vogel, R.M., Chapra, S.C., Ramsburg, C.A., 2014. Uncertainty and sensitivity analyses using glue when modeling inhibition and pharmaceutical cometabolism during nitrification. Environmental Modelling & Software 60, 219–227.

Shono, H., 2000. Efficiency of the finite correction of Akaike's information criteria. Fisheries Science 66, 608–610.

Shukor, M.Y.A., 2014. Evaluation of several mathematical models for fitting the growth of the algae dunaliella tertiolecta. Asian Journal of Plant Biology 2, 1–6.

Sittig, S., Kasteel, R., Groeneweg, J., Vereecken, H., 2012. Long-term sorption and sequestration dynamics of the antibiotic sulfadiazine: a batch study. Journal of Environmental Quality 41, 1497–1506.

Snip, L.J., Flores-Alsina, X., Plósz, B.G., Jeppsson, U., Gernaey, K.V., 2014. Modelling the occurrence, transport and fate of pharmaceuticals in wastewater systems. Environmental Modelling & Software 62, 112–127.

Ting, Y.F., Praveena, S.M., 2017. Sources, mechanisms, and fate of steroid

estrogens in wastewater treatment plants: a mini review. Environmental monitoring and assessment 189, 178.

Verlicchi, P., Al Aukidy, M., Zambello, E., 2012. Occurrence of pharmaceutical compounds in urban wastewater: removal, mass load and environmental risk after a secondary treatment—a review. Science of the Total Environment 429, 123–155.

Wilks, S.S., 1962. Mathematical statistics. a wiley publication in mathematical statistics john wiley & sons. Inc., New York-London .

Xu, Y., Yuan, Z., Ni, B.J., 2016. Biotransformation of pharmaceuticals by ammonia oxidizing bacteria in wastewater treatment processes. Science of the Total Environment 566, 796–805.

Yi, T., Harper, W.F., 2007. The effect of biomass characteristics on the partitioning and sorption hysteresis of 17α-ethinylestradiol. Water Research 41, 1543–1553.

Yi, T., Harper Jr, W.F., Holbrook, R.D., Love, N.G., 2006. Role of particle size and ammonium oxidation in removal of 17 α-ethinyl estradiol in bioreactors. Journal of Environmental Engineering 132, 1527–1529.

Yi, T., Mackintosh, S., Aga, D.S., Harper, W.F., 2011. Exploring 17α-ethinylestradiol removal, mineralization, and bioincorporation in engineered bioreactors. Water Research 45, 1369–1377.

Yu, W., Du, B., Yang, L., Zhang, Z., Yang, C., Yuan, S., Zhang, M., 2019. Occurrence, sorption, and transformation of free and conjugated natural steroid estrogens in the environment. Environmental Science and Pollution Research 26, 9443–9468.