Water Research

Formulation and probabilistic assessment of reversible biodegradation pathway of Diclofenac in groundwater

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research Paper
Keywords:	Diclofenac; denitrification; Reversible biodegradation; Bayesian calibration; Uncertainty quantification; Acceptance-Rejection Sampling
Corresponding Author:	Laura Ceresa, M.D. Politecnico di Milano Milano, ITALY
First Author:	Laura Ceresa
Order of Authors:	Laura Ceresa
	Alberto Guadagnini
	Giovanni Michele Porta
	Monica Riva
Abstract:	We present a conceptual and mathematical framework leading to the development of a biodegradation model capable to interpret the observed reversibility of the Pharmaceutical Sodium Diclofenac along its biological degradation pathway in groundwater. Diclofenac occurrence in water bodies poses major concerns due to its persistent (and bioactive) nature and its detection in surface waters and aquifer systems. Despite some evidences of its biodegradability at given reducing conditions, Diclofenac attenuation is often interpreted with models which are too streamlined, thus potentially hampering appropriate quantification of its fate. In this context, we propose a modeling framework based on the conceptualization of the molecular mechanisms of Diclofenac biodegradation which we then embed in a stochastic context, thus enabling one to quantify predictive uncertainty. We consider reference environmental conditions (biotic and denitrifying) associated with a set of batch experiments that evidence the occurrence of a reversible biotransformation pathway, a feature that is fully captured by our model. The latter is then calibrated in the context of a Bayesian modeling framework through an Acceptance-Rejection Sampling approach. By doing so, we quantify the uncertainty associated with model parameters and predicted Diclofenac concentrations. We discuss the probabilistic nature of uncertain model parameters and the challenges posed by their calibration with the available data. Our results are consistent with the recalcitrant behavior exhibited by Diclofenac in groundwater and documented through experimental data and support the observation that unbiased estimates of the hazard posed by Diclofenac to water resources should be assessed through a modeling strategy which fully embeds uncertainty quantification.
Suggested Reviewers:	Bill X. Hu Jinan University Institute of Groundwater and Earth Sciences 2776090374@qq.com Expertise in Stochastic Hydrogeology, Karst Hydrology and Seawater intrusion to groundwater.
	Xavier Sanchez-Vila Universitat Politecnica de Catalunya xavier.sanchez-vila@upc.edu Expertise in quantitative groundwater hydrology
	Ishai Dror Weizmann Institute of Science ishai.dror@weizmann.ac.il Expertise in Environmental Impact Assessment
	Tiziana Tosco Politecnico di Torino

Formulation and probabilistic assessment of reversible biodegradation pathway of Diclofenac in groundwater

Laura Ceresa, Alberto Guadagnini, Giovanni M. Porta, Monica Riva

Department of Civil and Environmental Engineering (DICA), Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano, Italy

Abstract

We present a conceptual and mathematical framework leading to the development of a biodegradation model capable to interpret the observed reversibility of the Pharmaceutical Sodium Diclofenac along its biological degradation pathway in groundwater. Diclofenac occurrence in water bodies poses major concerns due to its persistent (and bioactive) nature and its detection in surface waters and aquifer systems. Despite some evidences of its biodegradability at given reducing conditions, Diclofenac attenuation is often interpreted with models which are too streamlined, thus potentially hampering appropriate quantification of its fate. In this context, we propose a modeling framework based on the conceptualization of the molecular mechanisms of Diclofenac biodegradation which we then embed in a stochastic context, thus enabling one to quantify predictive uncertainty. We consider reference environmental conditions (biotic and denitrifying) associated with a set of batch experiments that evidence the occurrence of a reversible biotransformation pathway, a feature that is fully captured by our model. The latter is then calibrated in the context of a *Bayesian* modeling framework through an Acceptance-Rejection Sampling approach. By doing so, we quantify the uncertainty associated with model parameters and predicted Diclofenac concentrations. We discuss the probabilistic nature of uncertain model parameters and the challenges posed by their calibration with the available data.

*Corresponding author
Email address: laura.ceresa@polimi.it (Laura Ceresa)

Our results are consistent with the recalcitrant behavior exhibited by Diclofenac in groundwater and documented through experimental data and support the observation that unbiased estimates of the hazard posed by Diclofenac to water resources should be assessed through a modeling strategy which fully embeds uncertainty quantification.

Key words: Diclofenac, Denitrification, Reversible biodegradation, Bayesian calibration, Uncertainty quantification, Acceptance-Rejection Sampling

---

1. Introduction

Groundwater contamination by pharmaceuticals (PhAs) stands as a critical issue in modern society. Regulating authorities are recognizing potential risks associated with pollution of drinking water sources by biologically active compounds, such as PhAs, even as their detection is often limited to trace concentration levels ($1 - 1000[ng \cdot L^{-1}]$). Several classes of medical drugs are currently monitored according to international guidelines (CHMP, E.M.A. (2006)), with particular focus on the most persistent compounds. Sodium Diclofenac (NaDcf) ($NaDcf_{(s)}$, $C_{14}H_{10}Cl_2NO_2 \cdot Na_{(s)}$, CAS number 15307-79-6, $MW = 318.3[g \cdot mol^{-1}]$, Figure A.1 (a) in Appendix A) (Fonger et al. (2014)), a Non-Steroidal Anti-Inflammatory Drug (NSAID) is commonly prescribed as an analgesic (Small (1989)) and is seen to pose major concerns to surface waters and groundwater bodies (Lonappan et al. (2016)). Due to its suspected long-term impact on the ecology of aquatic environments upon chronic exposure, NaDcf has been classified by the Global Water Research Coalition (GWRC) as *Class 1* – high priority substance within relevant Pharmaceuticals to the water cycle (de Voogt et al. (2007)). Despite risk assessment protocols (CHMP, E.M.A. (2006)) often address Sodium Diclofenac biodegradation to non-hazardous compounds as a natural attenuation mechanism, several authors have recently questioned its effectiveness under reducing redox conditions, such as, e.g., denitrifying scenarios (Barbieri et al. (2012), Chiron and Duwig (2016)). While studies based on laboratory-scale batch systems (Barbieri et al.

(2011)), column experiments (Schaffer et al. (2015)) as well as field-scale data (Nham et al. (2015), Chiron and Duwig (2016)) are available in the literature, there is still a lack of knowledge about the behavior of Diclofenac (Dcf) in groundwater bodies, especially under scenarios comprising the occurrence of biologically mediated transformations. This observation suggests the need to further enhance *state-of-the-art* modeling frameworks and tools to support policies associated with environmental protection of aquatic compartments from risks related to bioaccumulation and biomagnification (CHMP, E.M.A. (2006)). In this context, it is noted that oversimplified biochemical models might lead to an inaccurate assessment of contaminant fate (Greskowiak et al. (2017)). Otherwise, relevant uncertainties might arise as a consequence of the complexity and non-linearity of model input-output relationships which can be associated with the representation of the richness of molecular dynamics associated with a full chemical reaction network that might in turn require a high level of model parametrization (Porta et al. (2018)). These uncertainties are hard to control and eventually constrain and reduce because of the paucity of available datasets documenting the fate of emerging contaminants and pharmaceuticals in groundwater. Therefore, implementation of an appropriate stochastic inverse modeling framework is key to address the effects of uncertainty propagation from the model structure and parameters to target model outputs (e.g., contaminant concentrations in environmental compartments), as a consequence of conditioning modeling results on available observations.

In this work, we focus on the fate of Diclofenac in groundwater under biotic, denitrifying redox conditions. These environmental conditions are relevant in field scale scenarios such as those that can take place across regions downstream of reactive barriers installed in the hyporheic zone. Such a scenario has been mimicked in a series of (biotic) *batch* experiments (Barbieri et al. (2012), their series $1[\mu g \cdot L^{-1}]$), which document the emergence of 5-C-Nitro-Diclofenac (NO_2Dcf) derivative ($NO_2Dcf_{(aq)}$), $C_{14}H_{10}Cl_2N_2O_{4(aq)}$, $MW = 341.15[g \cdot mol^{-1}]$, Figure A.2 (b) in Appendix A) and its reversible back-transformation into the parent compound once denitrifying conditions cease to occur. Several authors (Bar-

bieri et al. (2012), Chiron and Duwig (2016)) document the reversibility of Diclofenac transformation processes along its biologically mediated reaction pathway (under the above mentioned environmental conditions), which involves an almost complete back-transformation of its by-products into the parent compound. The latter process takes place upon complete Nitrite depletion, which is turn marked by the the occurrence of stronger reducing conditions than denitrification (such as, e.g., those corresponding to $Mn^{4+}_{(s)}/Mn^{2+}_{(aq)}$ in our scenario), and is supported by experimental observations (Barbieri et al. (2012), Chiron and Duwig (2016)).

Our study is keyed to the following three major objectives: (*i*) the development of a model capable of interpreting the observed reversible pathway of Diclofenac; (*ii*) the design of a workflow for the stochastic calibration of such a model through the estimation of its parameters against observations; and (*iii*) the interpretation of model parameters and outputs in a probabilistic sense, which fully embeds predictive uncertainty quantification. To achieve these objectives, we consider the set of previously mentioned microcosms experiments (Barbieri et al. (2011) and Barbieri et al. (2012)) that document the reversible behavior of the tranformation products (TPs) of Diclofenac and Sulphamethoxazole (SMX) downstream of the denitrification cycle, as discussed above. To interpret the selected dataset, we first formulate an original conceptual and biochemical model, which follows the same rationale employed in Rodríguez-Escales and Sanchez-Vila (2016) to model the fate of SMX, another amino compound used as antibiotic. Our stochastic inverse modeling approach relies on a *Bayesian* framework and rests on Acceptance-Rejection Sampling (ARS) technique to provide a probabilistic characterization of model parameters based on available data.

2. Methods

We first introduce the geochemical model we consider to interpret the documented reversible biodegradation pathway of Diclofenac. The key elements of

the implemented model are offered in Section 2.1. Additional details about the implementation strategy within the geochemical simulation Software used (pH RedOx Equilibrium C++ Software (PHREEQC), version 3.6.2, Parkhurst and Appelo (2013)) are provided in Supplementary Material S1. Further elements on the experimental framework of reference (Barbieri et al. (2012), Rodríguez-Escales and Sanchez-Vila (2016)) are available in the Supplementary Material S2. Section 2.2 describes the stochastic framework employed to calibrate the proposed model and characterize associated parameters and relevant outputs, including quantification of predictive uncertainty. Within the suite of *Bayesian* inverse modeling techniques, we rely on Acceptance-Rejection Sampling due to its unbiased nature. This approach yields the posterior (i.e., conditional on available data) probability distribution of uncertain model parameters within a *Bayesian* framework, circumventing the assumption of Gaussianity typical of, e.g., a Maximum Likelihood (ML) framework.

2.1. Geochemical model

We model the combined action of several mechanisms that can impact the temporal evolution of Diclofenac concentrations, i.e., precipitation/dissolution (see Supplementing Material S1.2.1), sorption/desorption to soil particles (see Supplementing Material S1.2.2), and biodegradation (a term which is here employed to embed several mechanisms that are discussed in details in Section 2.1.1). We illustrate here some details of the biochemical model employed to characterize biodegradation of Diclofenac. The full description of the model, including the complete collection of the considered chemical reactions is then provided in the Supplementary Materials and Appendices.

2.1.1. Reversible biodegradation pathway

Our model aims at interpreting the fate of Diclofenac at the environmental conditions investigated in Barbieri et al. (2012). We conceptualize the biochemical process according to the three Phases described in the following. These are schematically depicted in Figure 1 and further detailed in Figure 2.

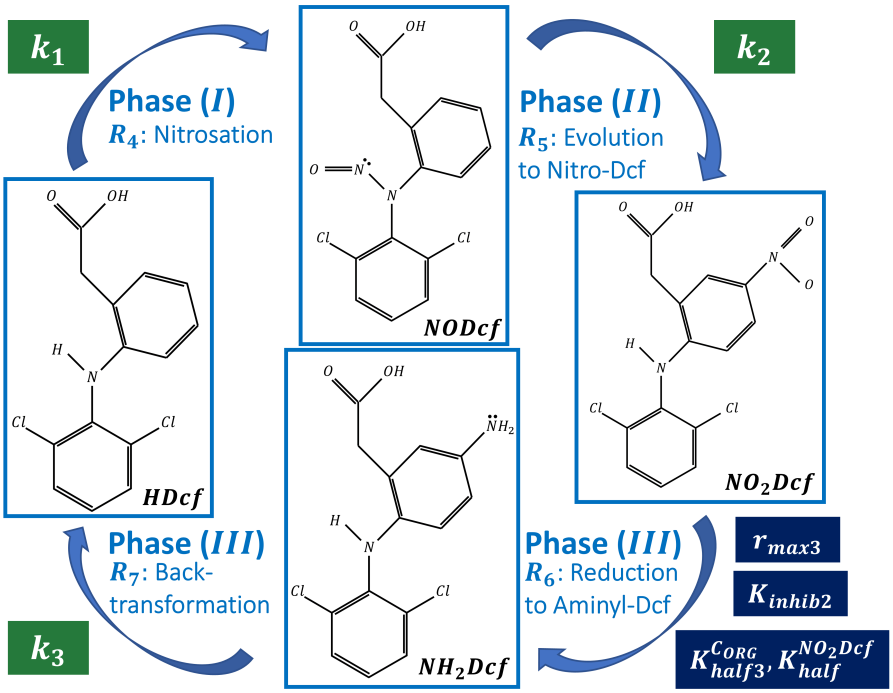
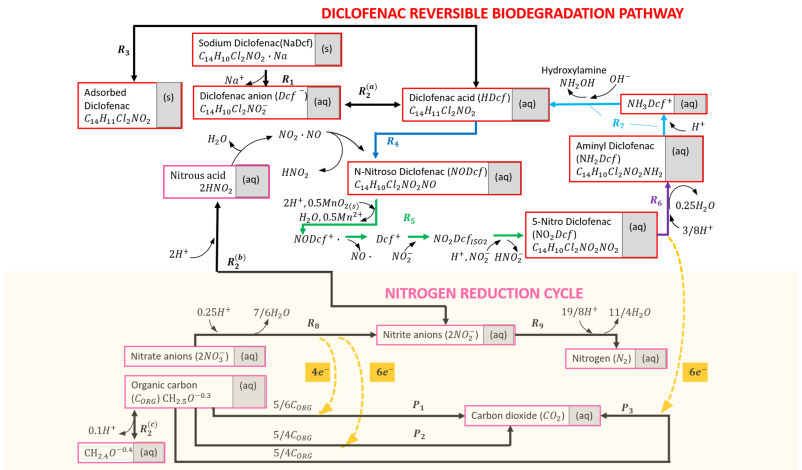


Figure 1: Reversible degradation pathway of Diclofenac under biotic, denitrifying redox conditions. Chemical processes leading to the release of transformation products are framed within specific phases, here denoted as Phase (I), (II), and (III). Main uncertain model parameters are also depicted within boxes (blue and green backgrounds of these denote whether the corresponding reactions are framed in the context of bacterial metabolism or not, respectively).



Equilibrium driven reactions:	*Kinetic reactions:*	*Kinetic biotransformation reactions (Michaelis-Menten-Monod):*
$\boldsymbol{R_1}$: Diclofenac precipitation / dissolution $\boldsymbol{R_2^{(i)}}$: Acid dissociation / aqueous complexation of *i-th* compound $\boldsymbol{R_3}$: Diclofenac adsorption due to surface complexation	$\boldsymbol{R_4}$: Diclofenac N-nitrosation by co-metabolism $\boldsymbol{R_5}$: Nitroso to Nitro-Diclofenac evolution $\boldsymbol{R_7}$: Electrophilic Aromatic Substitution	$\boldsymbol{R_6}$: Nitro to Aminyl-Diclofenac reduction by direct metabolism $\boldsymbol{R_8}$: Reduction N(V) to N(III) $\boldsymbol{R_9}$: Reduction N(III) to N(0) $\boldsymbol{P_1}$: Organic carbon C(-0.8) oxidation to C(IV): first path $\boldsymbol{P_2}$: Organic carbon C(-0.8) oxidation to C(IV): second path $\boldsymbol{P_3}$: Organic carbon C(-0.8) oxidation to C(IV): third path

Figure 2: Reaction network of Diclofenac reversible biodegradation.

- (Phase (I)) The abiotic degradation of $HDcf_{(aq)}$ to N-Nitroso-Diclofenac (NODcf) (i.e., $NODcf_{(aq)}$, $C_{14}H_{10}Cl_2N_2O_{3(aq)}$; see Figure A.2 (a) in Appendix A) takes place through *co-metabolism* under anoxic, biotic denitrifying redox conditions. Here, Diclofenac does not participate directly to the metabolic mechanism, which involves the (kinetically controlled) microbial reduction of Nitrate N(V) ($NO^-_{3(aq)}$) along the Nitrogen reduction cycle. This yields elemental Nitrogen N(0) ($N_{2(g)}$) upon complete conversion of Nitrite N(III)($NO^-_{2(aq)}$), a metastable intermediate product typically observed along the N(V) reduction cycle (Appelo and Postma (2004)). The process is sustained by the oxidation of dissolved organic matter (C_{ORG}) to Carbon dioxide C(IV) ($CO_{2(aq)}$). N-nitrosation of HDcf takes place together with aqueous complexation of Nitrite into Nitrous acid ($HNO_{2(aq)}$), a process that can be considered at instantaneous equilibrium and is characterized by pK_a of approximately 3.2 (Fonger et al. (2014)). The key role of Nitrous acid in this Phase is consistent with the guidelines of environmental geochemistry, which identify $HNO_{2(aq)}$ as a typical nitrosating agent at relevant (prevalently alkaline) environmental conditions (Stumm and Morgan (2012)). In this context, Diclofenac degradation is *abiotic* and *co-metabolic* because the compound does not participate directly to the redox metabolism of denitrification.

- (Phase (II)) The second Phase involves the evolution of NODcf to NO_2Dcf. We note that N-Nitroso-Diclofenac is typically considered an unstable intermediate product that rapidly evolves at appropriate reducing conditions (Smith (2020), Chiron and Duwig (2016)), such as those associated with, e.g., the presence of Pyrolusite ($MnO_{2(s)}$) acting as oxidizer. Here, Nitrites $NO^-_{2(aq)}$ simultaneously act as attacking reagents and Lewis base catalysts of the electrophilic aromatic substitution (EAS) yielding 5-C-Nitro-Diclofenac, which corresponds to the final TP experimentally detected by several authors (Barbieri et al. (2012), Chiron and Duwig (2016)) under denitrifying redox conditions. This process takes place

jointly with denitrification and is modeled as a multistage reaction. The involved steps are consistent with previous studies (Chiron and Duwig (2016)) and are here interpreted through a kinetic model. In this context, the evolution of NODcf is initiated by its one-electron oxidation sustained by Pyrolusite reduction. The latter yields N-Nitroso-$Dcf^{+\cdot}$ (Lewis formula $NODcf^{+\bullet}{}_{(aq)}$), a nitroso radical cation of NODcf which features a positive charge localized on the central Nitrogen of the amino functional group (Figure A.3 (a) in Appendix A). This is a highly unstable compound which undergoes a spontaneous fragmentation into Nitrous monoxide free radical (•NO) and Nitrenium cation, here denoted as Dcf^{+} (see Figure A.3 (b) in Appendix A). Consistent with the resonance theory (Smith (2020)), an instantaneous rearrangement of the positive charge is expected at this step (the shift occurring from the central Nitrogen of the amino functional group towards the topping ring). This yields the carbocation depicted in Figure A.3 (c) of Appendix A, a typical electrophile that can easily be neutralized by nucleophiles such as $NO_{2(aq)}^{-}$, consistent with the Lewis theory on acid-base reactions. Specifically, the reaction product is here a structural isomer (termed $NO_2Dcf_{ISO2(aq)}$; see Figure A.4 (a) of Appendix A) of the detected TP, 5-C-Nitro-Diclofenac. The last stage of Phase (*II*) involves instantaneous conversion of $NO_2Dcf_{ISO2(aq)}$ to the main isomer $NO_2Dcf_{(aq)}$. This occurs through simultaneous protonation of the double C=N bond of the amino group and one-proton removal in the *para* position of the topping ring, which takes place through the action of a strong Lewis base, possibly $NO_{2(aq)}^{-}$. The ring aromaticity is then restored, yielding the expected NO_2Dcf derivative (Figures A.2 (b) and A.4 (b) in Appendix A).

- (Phase (*III*)) The last Phase involves two consecutive reactions. First, we consider $NO_2Dcf_{(aq)}$ to undergo a reductive transformation to the corresponding amine 5-Aminyl-Diclofenac (NH_2Dcf) ($C_{14}H_{12}Cl_2N_2O_{(aq)}$, Figure A.2 (c) of Appendix A), consistent with previous studies on aquatic chemistry (Stumm and Morgan (2012)). Here, $NH_2Dcf_{(aq)}$ acts as a

meta-stable intermediate in the broader context of back-transformation processes to the parent compound. Consistent with Appelo and Postma (2004), a *direct* microbial transformation of 5-C-Nitro-Diclofenac takes place at this step, sustained by additional oxidation of organic Carbon. The second reaction of Phase (*III*) involves NH_2Dcf back-transformation to the parent compound through electrophilic aromatic substitution. In our approach we assume that dissolved protons $H^+_{(aq)}$ are responsible for protonation of NH_2Dcf, which is known to act as rate limiting step in the context of electrophilic aromatic substitutions. This yields a carbocation (termed $NH_3Dcf^+_{(aq)}$; see Figure A.5 (a) of Appendix A) whose aminyl group $NH^+_{2(aq)}$ is removed by the action of strong bases, possibly dissolved hydroxyl anions $OH^-_{(aq)}$. The final product is then the parent compound $HDcf_{(aq)}$, which is finally restored together with Hydroxylamine $NH_2OH_{(aq)}$ (Figure A.5 (b) in Appendix A), the secondary product of this reaction.

2.1.2. Quantitative model description

This Section is devoted to the illustration of the main details about the conceptual framework and the ensuing formulation of the mathematical model employed to interpret the reversible biochemical pathway introduced in Section 2.1.1. Consistent with the mixture used in the microcosms experiments of Barbieri et al. (2011) and with the formulations in Rodríguez-Escales and Sanchez-Vila (2016) (see the Supplementary material S1.1.1 for details), organic matter at Phase (*I*) is modeled as a mixture of Methanol C(-II) ($CH_4O_{(aq)}$) and Acetate anion C(0) ($C_2H_3O_2{}^-{}_{(aq)}$). For completeness, we recall that Rodríguez-Escales and Sanchez-Vila (2016) model the reaction rate of denitrification (per unit mole consumption of organic Carbon) according to the following *Multiplicative*

Monod equations (Appelo and Postma (2004)):

$$RR_{C_{ORG}}^{(REDOX1)}(t) = -\left.\frac{d\{\mathrm{C_{ORG}\,(aq)}\}(t)}{dt}\right|_{(REDOX1)}$$

$$= r_{max1}\frac{\{\mathrm{C_{ORG}\,(aq)}\}(t)}{\{\mathrm{C_{ORG}\,(aq)}\}(t) + K_{half1}^{\mathrm{C_{ORG}}}}\frac{\{\mathrm{NO_3^-\,(aq)}\}(t)}{\{\mathrm{NO_3^-\,(aq)}\}(t) + K_{half}^{\mathrm{NO_3^-}}}$$

$$\cdot\,\{\mathrm{CH_2O\,(s)}\} \tag{1}$$

$$RR_{C_{ORG}}^{(REDOX2)}(t) = -\left.\frac{d\{C_{ORG}\,(\mathrm{aq})\}(t)}{dt}\right|_{(REDOX2)} =$$

$$= r_{max2}\frac{\{\mathrm{C_{ORG}\,(aq)}\}(t)}{\{\mathrm{C_{ORG}\,(aq)}\}(t) + K_{half2}^{\mathrm{C_{ORG}}}}\frac{\{\mathrm{NO_2^-\,(aq)}\}(t)}{\{\mathrm{NO_2^-\,(aq)}\}(t) + K_{half}^{\mathrm{NO_2^-}}}$$

$$\cdot\,\frac{K_{inhib}}{\{\mathrm{NO_3^-\,(aq)}\}(t) + K_{inhib}}\{\mathrm{CH_2O\,(s)}\} \tag{2}$$

where $\mathrm{CH_2O\,(s)}$is the biomass molarity, $r_{max_i},\ i = 1, 2$ is the maximum rate of substrate consumption relative to biomass in the $i - th$ redox reaction, $K_{half_i},\ i = \mathrm{C_{ORG1}}, \mathrm{C_{ORG2}}, \mathrm{NO_3^-}, \mathrm{NO_2^-}$ are half saturation constants, and the inhibition constant K_{inhib} embeds the hindering effect of Nitrates on Nitrites reduction. The notation $\{\cdot\}$ identifies species activity, that tends to coincide with molar concentration in very diluted solutions (Appelo and Postma (2004)). Note that these parameters are not subject to calibration in our work, being rather fixed to the values estimated by Rodríguez-Escales and Sanchez-Vila (2016) (Table 2 in Section 2.2).

Diclofenac nitrosation takes place together with aqueous complexation into Nitrous acid, according to the following stages (Mirvish (1975)): (a) first, two moles of Nitrous acid rapidly dissociate in the nitrosyl carrier $\mathrm{NO_2{\bullet}NO}$ (Nitrous anhydride) and water; then, (b) $\mathrm{NO_2{\bullet}NO}$ reacts with the secondary amine, yielding N-Nitroso-Diclofenac (rate limiting step):

$$2\,\mathrm{HNO_{2(aq)}} \overset{\mathrm{Keq_1}}{\rightleftharpoons} \mathrm{NO_2{\bullet}NO_{(aq)}} + \mathrm{H_2O} \tag{3}$$

$$\mathrm{HDcf_{(aq)}} + \mathrm{NO_2{\bullet}NO_{(aq)}} \xrightarrow{\mathrm{k}} \mathrm{NODcf_{(aq)}} + \mathrm{HNO_{2(aq)}} \tag{4}$$

The global stoichiometry is:

$$\mathrm{HNO_{2(aq)}} + \mathrm{HDcf_{(aq)}} \xrightarrow{\mathrm{k_1}} \mathrm{NODcf_{(aq)}} + \mathrm{H_2O} \tag{5}$$

Diclofenac nitrosation is therefore characterized by a reaction rate of order two with respect to Nitrous acid concentration and of order three globally, i.e.,:

$$\begin{aligned} RR_{NODcf}(t) &= \frac{d\{\mathrm{NODcf(aq)}\}(t)}{dt} = k\{\mathrm{HDcf(aq)}\}(t)\{\mathrm{NO_2{}^{\bullet}NO}\} \\ &= k\{\mathrm{HDcf(aq)}\}(t)K_{eq1}\{\mathrm{HNO_2(aq)}\}^2(t) \\ &= k_1\{\mathrm{HDcf(aq)}\}(t)\{\mathrm{HNO_2(aq)}\}^2(t) \end{aligned} \tag{6}$$

where $k_1 = kK_{eq1}$ is a rate constant. For simplicity, we neglect here the temporal evolution of biomass concentration, which is set to the same value as in Rodríguez-Escales and Sanchez-Vila (2016).

Phase (*II*) is modeled assuming that one-electron oxidation of HDcf is governed by redox equilibria, the reductive half-reaction involving the redox couple $\mathrm{MnO_{2(s)}/Mn^{+}_{2(aq)}}$. The spontaneous fragmentation yielding Nitrenium cations is then considered instantaneous, consistent with the high instability of the latter. The nucleophilic attack of Nitrite to the carbocation is then rate limiting, in agreement with the chemistry of Lewis acid-base reactions (Smith (2020)) and also consistent with the presence of a substituted amino group ($-\mathrm{NHC_6H_3Cl_2}$) on the top (aromatic) ring, which is *para* activating towards further substitution (on the very same ring). Lastly, the central Nitrogen protonation in the amino group (taking place after neutralization of the positive charge by $\mathrm{NO^{-}_{2(aq)}}$ attack to the carbocation) is again instantaneous.

These steps can be expressed as:

$$\mathrm{NODcf_{(aq)} + 0.5\,MnO_{2(s)} + 2\,H^{+}_{(aq)} \overset{K_{redox}}{\rightleftharpoons} NODcf^{+\bullet}{}_{(aq)} + 0.5\,Mn^{2+}_{(aq)} + H_2O} \tag{7}$$

$$\mathrm{NODcf^{+\bullet}{}_{(aq)} \xrightarrow{instantaneous} {}^{\bullet}NO_{(aq)} + Dcf^{+}_{(aq)}} \tag{8}$$

$$\mathrm{Dcf^{+}_{(aq)} + NO_2{}^{-}{}_{(aq)} \xrightarrow{k} NO_2Dcf_{ISO2(aq)}} \quad \text{(secondary structural isomer)} \tag{9}$$

$$\mathrm{NO_2Dcf_{ISO2(aq)} + H^{+}_{(aq)} + NO_2{}^{-}{}_{(aq)} \xrightarrow{fast} NO_2Dcf_{(aq)} + HNO_{2(aq)}}$$

$$\text{(main structural isomer)} \tag{10}$$

Thus, the global stoichiometry becomes:

$$\mathrm{NODcf_{(aq)} + 0.5\,MnO_{2(s)} + 3\,H^{+}_{(aq)} + 2\,{NO_2}^{-}{}_{(aq)} \xrightarrow{k_2}} \\ \mathrm{0.5\,Mn^{2+}_{(aq)} + {\bullet}NO_{(aq)} + H_2O + NO_2Dcf_{(aq)} + HNO_{2(aq)}} \tag{11}$$

The global reaction rate can be formulated as:

$$RR_{NO_2Dcf}(t) = \frac{d\{\mathrm{NO_2Dcf\,(aq)}\}(t)}{dt} = k\{\mathrm{NO_2^-\,(aq)}\}(t)\{\mathrm{Dcf^+\,(aq)}\}(t) \\ = k_2 \frac{\{\mathrm{NO_2^-\,(aq)}\}(t)\{\mathrm{H^+\,(aq)}\}^2(t)\{\mathrm{NODcf\,(aq)}\}(t)}{\sqrt{\{\mathrm{Mn^{2+}\,(aq)}\}(t)}} \tag{12}$$

where $k_2 = k \cdot K_{redox}$ is the process rate constant.

The back-transformation into the parent compound (Phase (*III*)) entails two consecutive reactions. The first one involves the direct metabolic transformation of NO_2Dcf into NH_2Dcf (corresponding to microbial reduction sustained by organic Carbon oxidation) whose kinetics are modeled according to the *Michaelis-Menten-Monod* mathematical framework (Appelo and Postma (2004)) (specifically, through the *Multiplicative Monod* equations). The global stoichiometry (see Supplementary Material S1.1.1) can be expressed as:

$$\mathrm{CH_4O_{(aq)} + \frac{3}{4}C_2H_3O_2{}^{-}{}_{(aq)} + \frac{3}{4}H^{+}_{(aq)} + 2NO_2Dcf_{(aq)} \xrightarrow{r_{max3}}} \\ \mathrm{\frac{5}{2}CO_{2(aq)} + 2NH_2Dcf_{(aq)} + \frac{1}{2}H_2O_{(aq)}} \tag{13}$$

This reaction involves reduction of the nitro group to aromatic amine through a six-electron transfer mechanism, in agreement with Razo-Flores et al. (1997) and Kulkarni and Chaudhari (2007). Note that the Methanol/Acetate mixture representing organic matter can be modeled through a single molecule of organic Carbon (C_{ORG}), the latter being implemented in PHREEQC as $CH_{2.5}O^{-0.3}$ (further details are then available in Supplementary Material S1.1.1).

The second stage involves the EAS of the aminyl group in NH_2Dcf. Here, the protonation step is rate limiting, in agreement with Smith (2020). The subsequent withdrawal of aminyl groups by dissolved hydroxyl anions is typically

very fast and is then considered to take place instantaneously.

This second reaction is then described by the following stages:

$$\mathrm{NH_2Dcf_{(aq)}} + \mathrm{H^+_{(aq)}} \xrightarrow{\mathrm{k_3}} \mathrm{NH_3Dcf^+_{(aq)}} \tag{14}$$

$$\mathrm{NH_3Dcf^+_{(aq)}} + \mathrm{OH^-_{(aq)}} \xrightarrow{\mathrm{fast}} \mathrm{HDcf_{(aq)}} + \mathrm{NH_2OH_{(aq)}} \tag{15}$$

This leads to the following global stoichiometry:

$$\mathrm{NH_2Dcf_{(aq)}} + \mathrm{H_2O} \xrightarrow{\mathrm{k_3}} \mathrm{HDcf_{(aq)}} + \mathrm{NH_2OH_{(aq)}} \tag{16}$$

where $\mathrm{NH_2OH_{(aq)}}$ (Figure A.5 in Appendix A) represents Hydroxylamine, which is another TP for which experimental investigations are still very scarce for the purposes of our study (i.e., it has not been monitored in the experiments of Barbieri et al. (2012)).

The corresponding reaction rates are rendered as:

$$\begin{aligned} RR_{NH_2Dcf}(t) &= \frac{d\{\mathrm{NH_2Dcf\,(aq)}\}(t)}{dt} \\ &= r_{max3}\frac{\{\mathrm{C_{ORG}\,(aq)}\}(t)}{\{\mathrm{C_{ORG}\,(aq)}\}(t) + K^{\mathrm{C_{ORG}}}_{half3}}\frac{\{\mathrm{NO_2Dcf\,(aq)}\}(t)}{\{\mathrm{NO_2Dcf\,(aq)}\}(t) + K^{\mathrm{NO_2Dcf}}_{half}} \\ &\quad \frac{K_{inhib2}}{\{\mathrm{NO_2^-\,(aq)}\}(t) + K_{inhib2}}\{\mathrm{CH_2O\,(s)}\} \end{aligned} \tag{17}$$

$$\begin{aligned} RR_{BT}(t) &= -\frac{d\{\mathrm{NH_2Dcf\,(aq)}\}(t)}{dt} = \frac{d\{\mathrm{HDcf\,(aq)}\}(t)}{dt} \\ &= k_3\{\mathrm{H^+\,(aq)}\}\{\mathrm{NH_2Dcf\,(aq)}\} \end{aligned} \tag{18}$$

Here, r_{max3} is the maximum *Michaelis-Menten-Monod* rate; $K^{\mathrm{C_{ORG}}}_{half3}$, $K^{\mathrm{NO_2Dcf}}_{half}$, and k_{inhib2} are half-saturation and inhibition constants; and k_3 is the rate constant of the last reaction of Phase (*III*). Note that equation (17) considers inhibition on the reduction of nitroaromatics in the presence of higher-priority oxidizers, such as Nitrites, which are associated with lower standard potentials of reduction (see, e.g., Appelo and Postma (2004)).

The chemical model is completed by three additional processes, which are considered at chemical equilibrium: (*i*) Diclofenac acid adsorption to soil particles, (*ii*) dissolution of Sodium Diclofenac in water, and (*iii*) dissociation of all of the (relevant) acidic compounds present in solution (i.e., Nitric and Nitrous acids,

Carbonic acid, HDcf and organic matter, as partially derived from the Acetic acid). Sorption is modeled as solely due to surface complexation of the neutral Diclofenac acid onto the organic Carbon fraction of the soil, the anionic form being insensitive to adsorptive mechanisms at the environmental conditions considered (Appelo and Postma (2004)). Surface complexation of HDcf is associated with a linear distribution coefficient (consistent with the very diluted conditions of the aqueous solution analyzed), defined as:

$$k_d = \frac{\{\mathrm{HDcf}_{(ADSORBED)}\}}{\{\mathrm{HDcf}_{(\mathrm{aq})}\}} \tag{19}$$

Considering a linear isotherm approach and following the procedure detailed in the Supplementary Material S1.2.1, we obtain $k_d \approx 3.6 \cdot 10^{-3}[L/kg_{sed}]$, clearly evidencing the negligible impact of sorption to the extent of Diclofenac degradation (additional details are provided in the above mentioned Supplementary Material).

The dissolution of Sodium Diclofenac is then modeled as:

$$\mathrm{NaDcf_{(s)}} \downarrow \underset{^{1}/_{k_{eq}}}{\overset{k_{eq}}{\rightleftharpoons}} \mathrm{Na^{+}_{(aq)}} + \mathrm{Dcf^{-}_{(aq)}} \tag{20}$$

$$k_{eq} = \frac{\{\mathrm{Na^{+}\,(aq)}\}\{\mathrm{Dcf^{-}\,(aq)}\}}{\{\mathrm{NaDcf\,(s)}\}} = \frac{k_s}{\cancel{\{\mathrm{NaDcf\,(s)}\}}} = k_s \tag{21}$$

where $k_s = 2.3 \cdot 10^{-10}[mol^2 \cdot L^{-2}]$ represents the solubility constant of the salt, as inferred from available databases (Wishart et al. (2006)). Further details are provided in Supplementary Material S1.2.1, where we show that one could consider in principle the full amount of NaDcf as fully dissolved in solution, consistent with the value of the solubility product and the limited (trace) concentration levels at which Sodium Diclofenac is typically detected in groundwater.

We finally note that acid dissociation processes govern the degree of ionization of the molecules (water speciation). In this context, dissociation of Diclofenac acid and organic matter (the same rationale applying for other molecules of

interest) can be formulated as (Wishart et al. (2006)):

$$\mathrm{HDcf_{(aq)} \underset{^{1}/_{\mathrm{Ka}}}{\overset{\mathrm{Ka}}{\rightleftharpoons}} H^{+}_{(aq)} + Dcf^{-}_{(aq)}} \tag{22}$$

$$\mathrm{p}K_{\mathrm{a}} = -Log_{10}K_{\mathrm{a}} = -Log_{10}\left\{ \frac{\{\mathrm{H^{+}_{(aq)}}\}\{\mathrm{Dcf^{-}_{(aq)}}\}}{\{\mathrm{HDcf_{(aq)}}\}} \right\} \approx 4.2 \tag{23}$$

$$\frac{3}{10}\mathrm{C_2H_3O_2{}^{-}{}_{(aq)}} + \frac{2}{5}\mathrm{CH_4O_{(aq)}} \underset{^{1}/_{\mathrm{Ka}}}{\overset{\mathrm{Ka}}{\rightleftharpoons}} \frac{3}{10}\mathrm{C_2H_4O_{2(aq)}} + \frac{2}{5}\mathrm{CH_3O^{-}_{(aq)}}\frac{1}{10}\mathrm{H^{+}_{(aq)}} \tag{24}$$

$$\mathrm{p}K_{\mathrm{a}} = -Log_{10}K_{\mathrm{a}} = \frac{\{\mathrm{C_2H_4O_{2(aq)}}\}^{0.3} \cdot \{\mathrm{CH_3O^{-}_{(aq)}}\}^{0.4} \cdot \{\mathrm{H^{+}_{(aq)}}\}^{0.1}}{\{\mathrm{C_2H_3O_2{}^{-}{}_{(aq)}}\}^{0.3} + \{\mathrm{CH_4O_{(aq)}}\}^{0.4}} = -4.772 \tag{25}$$

Our implementation fully includes also this latter process, additional details being available in Supplementary Materials S1.2.1 and S1.2.3 for Dcf and organic matter, respectively.

We also note that in our study the temporal evolution of Diclofenac concentrations always refers to the *total* molarity of Dcf *master* species (Parkhurst and Appelo (2013)), i.e., the overall contribution of its anionic and undissociated forms.

2.2. Stochastic model calibration

The mathematical representation of the biochemical setting we consider is rendered through the system of equations (1), (2), (6), (12), (17), (18), (19), (21), (23) and (25). This formulation is here applied under the prescribed initial conditions illustrated in the Supplementary Material S2 to interpret the data listed in Table 3. Stochastic model calibration is here performed upon considering the seven model parameters listed in Table 1 as uncertain, the remaining model parameters (listed in Table 2) being assumed as known, on the basis of prior studies (Rodríguez-Escales and Sanchez-Vila (2016)).

Our uncertainty analysis is framed in the context of *Bayesian* model calibration, which is performed through Acceptance-Rejection Sampling. In the absence of additional information about possible ranges of variability of the model parameters, the latter are obtained starting from a preliminary fit against the available

data. This step involves obtaining preliminary estimates of uncertain model parameters through a satisfactory visual agreement between experimental data and simulation results. The parameter support is then assessed upon considering two logarithmic cycles centered around these preliminary estimates. The uncertain parameters are then considered as independent and identically distributed (*iid*) random variables, each characterized by a uniform distribution within the intervals listed in Table 1. The choice of the latter distribution enables one to give the same weight to all parameter values across their support. Random sampling of parameter values within the corresponding support is performed through a classical Quasi-Monte Carlo technique (Sobol (1998)). Practical

Uncertain parameter	Lower limit	Upper limit
$k_1[\frac{L^2}{mol^2 s}]$	$1.2 \cdot 10^8$	$1.2 \cdot 10^{10}$
$k_2[\frac{L^{0.4}}{mol^{0.4} s}]$	$1.3 \cdot 10^2$	$1.3 \cdot 10^4$
$r_{max3}[\frac{1}{s}]$	$5.0 \cdot 10^{-12}$	$5.0 \cdot 10^{-10}$
$K_{half3}^{\mathrm{C_{ORG}}}[M]$	$1.0 \cdot 10^{-7}$	$1.0 \cdot 10^{-5}$
$K_{half}^{\mathrm{NO_2Dcf}}[M]$	$7.0 \cdot 10^{-10}$	$7.0 \cdot 10^{-8}$
$K_{inhib2}[M]$	$5.0 \cdot 10^{-7}$	$5.0 \cdot 10^{-5}$
$k_3[\frac{L}{mol \cdot s}]$	$5.0 \cdot 10^3$	$5.0 \cdot 10^4$

Table 1: Intervals of variability considered for the uncertain model parameters in the context of stochastic model calibration.

implementation of the ARS algorithm (Bolstad and Curran (2016)) for model parameter estimation relies upon embedding multiple model simulations based on the widely known and tested software PHREEQC (Parkhurst and Appelo (2013)) within the procedure outlined in the following. Acceptance-Rejection Sampling is designed to draw samples from the posterior density ($f_{\boldsymbol{p}|\boldsymbol{C}}$) of the considered random parameter set, conditional on observations of a given target quantity. Such a density is proportional to the likelihood function $f_{\boldsymbol{C}|\boldsymbol{p}}$, which is taken to be multi-Gaussian, available data being associated with independent

Parameter	Value
r_{max1} $[\frac{\mathrm{mM}}{\mathrm{d}}]$	19
r_{max2} $[\frac{\mathrm{mM}}{\mathrm{d}}]$	11
$K_{half1}^{\mathrm{C_{ORG}}}$ $[M]$	$1.6 \cdot 10^{-1}$
$K_{half2}^{\mathrm{C_{ORG}}}$ $[M]$	$1.8 \cdot 10^{-2}$
$K_{half}^{\mathrm{NO_3^-}}$ $[M]$	$1.0 \cdot 10^{-4}$
$K_{half}^{\mathrm{NO_2^-}}$ $[M]$	$5.0 \cdot 10^{-4}$
K_{inhib} $[M]$	$1.0 \cdot 10^{-4}$
$\{\mathrm{CH_2O\,(s)}\}$ $[mM]$	1

Table 2: Model parameters which are considered as fixed from prior studies (Rodríguez-Escales and Sanchez-Vila (2016)).

	C^*			
$t[d]$	$\frac{\{\mathrm{Dcf}\}}{\{\mathrm{Dcf}\}_0}$	$\frac{\{C_{ORG}\}}{\{C_{ORG0}\}}$	$\frac{\{\mathrm{NO_2^-}\}}{\{\mathrm{N}(V)_0\}}$	$\frac{\{\mathrm{NO_3^-}\}}{\{\mathrm{N}(V)_0\}}$
1.55	94.5%	–	–	–
3	70%	57%	49%	45%
5	56%	34%	63%	11%
10	91%	17%	0%	0%
20	83%	–	–	–

Table 3: Available observations for chemical species concentrations. Values are normalized by the initial concentration of the corresponding *master* species ({Dcf} denotes total Diclofenac concentration, including both its anionic and undissociated forms.)

errors characterized by a variance σ^2_{OBS}.

ARS enables one to draw random samples from $f_{\boldsymbol{p}|\boldsymbol{C}}$ according to the following workflow:

(1) Sample nmc random combinations of the seven uncertain model parameters from the corresponding uniform priors whose supports are listed in Table 1;

(2) Evaluate the temporal evolution of the concentrations of the compounds of interest (i.e., Dcf, organic Carbon, Nitrate and Nitrite) through the geochemical model for each of the nmc parameter combinations;

(3) Compute nmc acceptance probabilities α_i, $i = 1, 2, ..., nmc$ according to:

$$\alpha_i = \frac{f_{\boldsymbol{C}|\boldsymbol{p}_i}}{\max\left(f_{\boldsymbol{C}|\boldsymbol{p}_i}\right)} = \exp\left(-\frac{1}{2\sigma^2_{OBS}}\left[\boldsymbol{C}_i^{*} - \boldsymbol{C}_i\right]^T\left[\boldsymbol{C}_i^{*} - \boldsymbol{C}_i\right]\right) \tag{26}$$

where N^* is the total number of available observations; $C_i(t)$ is the model-based concentration of species i at time t, normalized by its initial concentration; $\mathbf{p}$ is a vector whose entries correspond to the (seven) uncertain model parameters, and C_i^* are the observed data.

(4) Draw nmc random values u_i ($i = 1, 2, ..., nmc$) from a uniform *pdf* in the unit support;

(5) Accept a given model realization i if the corresponding sampled value of u_i is smaller than α_i.

The number of Monte Carlo simulations should be designed to guarantee that stable posterior *pdf*s of model parameters are obtained. We note that the rate of acceptance associated with the algorithm depends on σ^2_{OBS}, whose precise assessment is clearly affected by quality and quantity of available information. While large values of σ^2_{OBS} typically lead to a large number of accepted realizations, these are also related to lower data quality. Since no precise information about measurement uncertainty is available for the data-set analyzed, we select σ^2_{OBS} as a generally reasonable compromise between a good acceptance rate and the loss of quality of the data (see also Section 3).

3. Results and Discussion

We present here the results of our study and provide a quantitative analysis of our modeling framework for the characterization of Diclofenac reversible biodegradation pathway.

Following the procedure detailed in Section 2.2, Acceptance-Rejection Sampling is performed upon setting $\sigma^2_{OBS} \approx 0.018$, as result of a compromise between achieving a good acceptance rate and considering data which are not associated with marked loss of quality, the total number of realizations considered in the ARS approach being set to $nmc = 200000$, yielding a number of accepted realizations in the order of 10^2 (details not shown). Figure 3 depicts the marginal distributions obtained for the seven considered model parameters conditional on the available observations. For completeness, the prior (uniform) densities (see Table 1) associated with each model parameters are juxtaposed to the ARS-based inverse modeling results. The resulting posterior means and intervals (centered around the mean) of width equal to a standard deviation are highlighted through vertical, dashed lines. Analysis of these results suggests that the posterior *pdf* of k_1 is the one which is most affected by conditioning as it markedly differs from its corresponding prior *pdf*. Conditioning on data is seen to affect also the *pdf*s of k_2 and k_3. All of these posterior densities display well defined peaks, corresponding to the Maximum A Posteriori (MAP) estimate, the latter being equal to the mode of the posterior distribution (identified by vertical dashed blue lines in Figure 3). While the posterior marginals of k_1 and k_2 are nearly symmetric, with modes very close to the corresponding mean values, the distribution of k_3 is visibly left-skewed, conditioning on data favoring the largest values of this parameter. As a consequence, stochastic calibration indicates that the back transformation steps (in Phase (*III*)) tends to take place with the fastest rates among those analyzed. These features are reinforced by the analysis of the corresponding box-plots depicted in Figure 4. The latter results yield a visual appraisal of the width of credible intervals of the model parameters, here associated with one inter-quartile range (i.e., in-

cluded within the first and the third quartiles) of the distributions of k_1, k_2, and k_3 and the relative location of the Maximum A Posteriori estimates within the corresponding prior supports. Figures 3 and 4 evidence the significant reduction of the support of the posterior distribution of k_1 with respect to the corresponding assumed prior. This is consistent with the observation that conditioning on data through (stochastic) inverse modeling yields a reduction of the uncertainty which is associated with all relevant model parameters in the absence of measurements. For example, one can note that conditioning on data leads to a reduction of the prior variance of k_1 by approximately 88%, thus providing a quantitative metric to support the strong influence of conditioning on the assessment of this parameter. Variances of k_2 and k_3 are seen to drop by about 34 and 8%, respectively, as compared against their prior counterparts. Otherwise, prior and posterior distributions of the remaining four model parameters display negligible differences. This is indicative of how no particular value of these parameters can be identified as most likely to interpret the available data. The latter is typically considered as an indication that such parameters are not influential to the overall variability of the model output following conditioning to the available observation. As such, any of the values of these parameters comprised within their range of variability is characterized by the same likelihood of being consistent with the data considered. Therefore, one should not expect significant uncertainty reduction for these parameters following acquisition of the type of data here considered. Notice that three of these parameters are associated with half-saturation and inhibition constants appearing in *Michaelis-Menten-Monod* kinetics. This result is consistent with the fact that observed concentrations attain low values, thus explaining why half-saturation and inhibition constants do not play a major role on the kinetic processes represented by reaction (13). As an example of the quality of the results, Figure 5 depicts outputs associated with the proposed geochemical model. Here, consistent with the stochastic model calibration framework, k_1, k_2, and k_3 are set at their MAP, while the remaining model parameters are characterized through their corresponding average value. These results are complemented

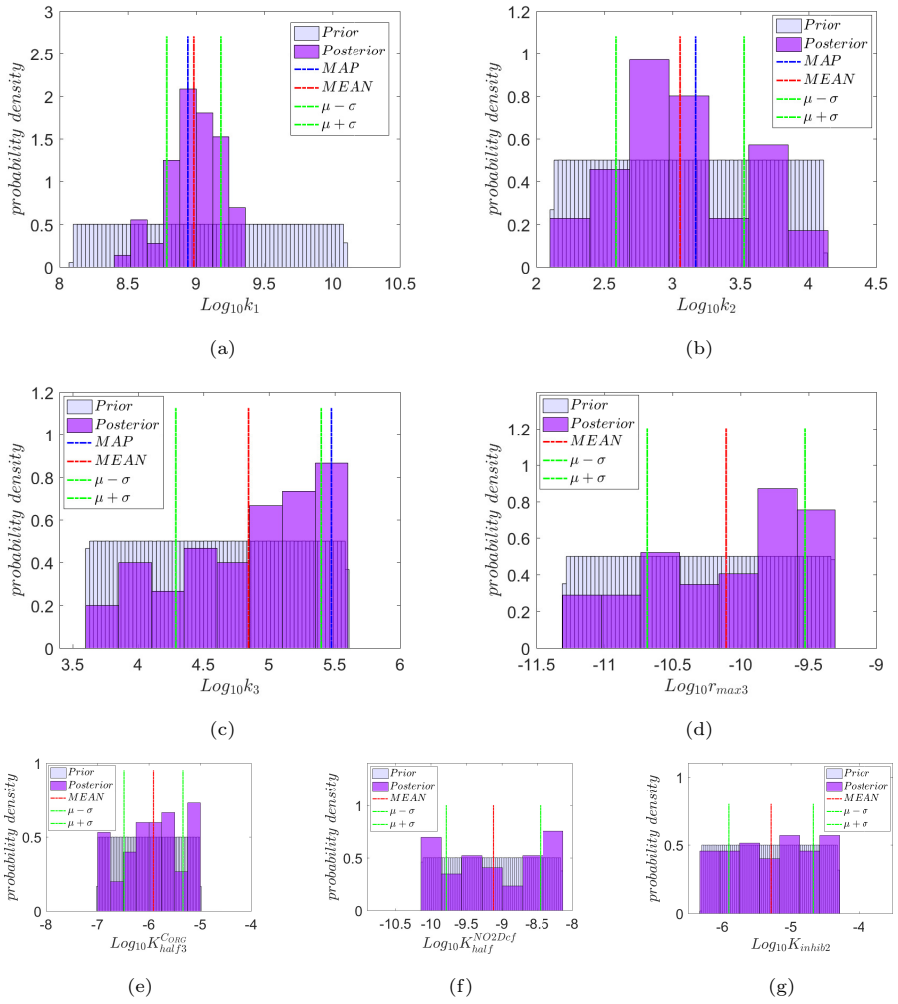


Figure 3: Prior and posterior *pdf*s of the model parameters, together with MAP estimates (dashed blue lines), mean values (dashed red lines). Dashed green lines delineate intervals of width equal to twice the standard deviation and centered around the corresponding mean value.

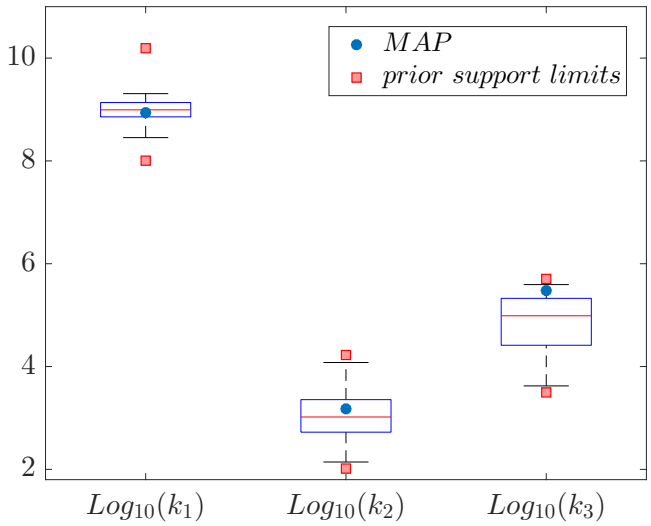


Figure 4: Boxplot representation of the marginal distributions of parameter values resulting from the inverse modeling procedure. Blue circles correspond to MAP estimates, red squares denoting lower and upper limits of the support of the prior distributions.

by Table 4 where we list parameter values associated with modeling results of Figure 5. One can see that considering MAP estimates of model parameters is conducive to high quality estimates of the fate of Diclofenac, even as it is evident that the MAP-based geochemical model tends to slightly underestimate the reaction progresses along the depicted temporal window. A good agreement is still observed amongst model- and observation-based concentrations for most of the main compounds undergoing redox reactions along the Nitrogen reduction path (Figure 5a). A non-negligible discrepancy is otherwise observed with reference to the temporal trend of organic Carbon degradation, our modeling results clearly underestimating the associated consumption rate. The latter result is consistent with (a) the observation that the temporal evolution of biomass concentration in the system has been neglected due to the limited dataset available, which prevents reliance on a more complex modeling approach; and (b) the study of Barbieri et al. (2012) who derive similar conclusions about the fate of organic matter along the biodegradation pathway. Indeed, these authors highlight that a relevant percentage (around 27%) of organic Carbon is possi-

bly consumed due to the action of additional processes, which are not included in their geochemical model. On these bases, and in line with Barbieri et al. (2012), an additional contribution of about 2.2$[mM]$ to the net consumption of organic Carbon is expected due to its further degradation as substrate that sustains bacterial growth (additional details on associated biochemistry are offered in the Supplementary Material S1.1.2).

As an example of the benefits arising from relying on a stochastic model calibration, Figure 6 shows the results of uncertainty propagation to the Dcf concentration history. Figure 6a depicts the temporal behavior of selected percentiles (i.e., 5^{th}, 50^{th}, and 95^{th}) as wells as the mean of the probability distribution of (normalized) Dcf concentrations resulting from our modeling study. Experimental observations are also depicted as a reference, together with the temporal evolution of Dcf concentrations obtained through the collection of some exemplary model realizations associated with the posterior probability densities of model parameters (solid grey curves). Figure 6b completes the picture upon showing *pdf*s of Dcf concentrations conditional on available data at three selected observation times. The width of the intervals associated with values comprised between the 5^{th} and 95^{th} percentile of the distribution widely vary across the temporal window considered, attaining a seemingly stable value at late time. The latter feature is consistent with our expectations, as our model does not implement any additional process downstream of the back-transformation to the parent compound, which is expected to end significantly sooner than the last monitored time (20 days).

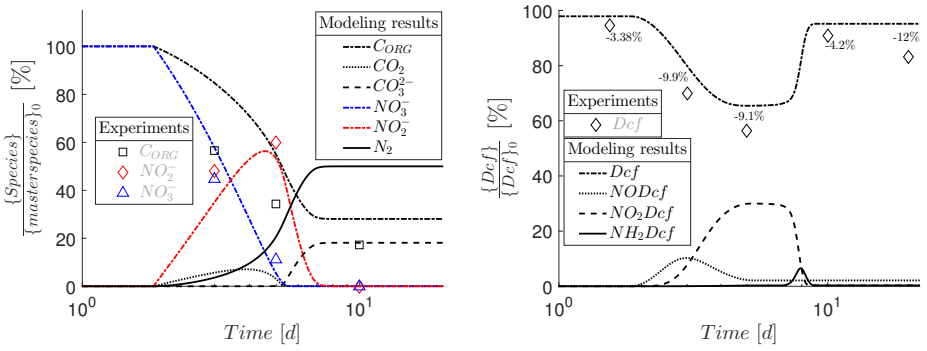
The available measurements are well within the intervals delineated by the (posterior) 5^{th} and 95^{th} percentiles. An exception is noted at early times (i.e., $t = 3$ days), with reference to a possible under-estimation of the nitrosation rate of HDcf, consistent with the trend outlined in Figure 5. This might be indicative of faster nitrosation rates than those associated with our model, whose results are otherwise fully consistent with the other available data. Finally, we observe that the initial decrease and successive increase of Dcf concentration tend to attain a similar rate, i.e. concentration histories attain a roughly symmetric U-

shape around time $5-6$ days. In particular, the steep increase of Dcf observed at time $7-8$ days for certain realizations (Figure 6a) can be linked with the relatively high probability of observing large values of the back-transformation rate constant k_3 (see Figure 3).

Our probabilistic results can then assist to assess the uncertainty linked to model outputs of interest also at unsampled times. In this context, exemplary probability density functions of Dcf are depicted in Figure 6c at unsampled times. The ability to have at our disposal these types of results is critical to support, for example, probabilistic risk assessment under uncertainty where one can be interested in quantifying the probability of exceeding a given threshold concentration in time. For example, these results suggest that posterior densities tend to almost overlap at sufficiently late times (such as, e.g., $t = 18$ and $t = 21$ days). Notice that this behavior is consistent with our expectations, as Diclofenac concentrations are not expected to undergo further variations after completion of back-transformation processes (denoted as Phase (III) in Section 2.1.1).

Parameter type	Parameter value	Log_{10} of Parameter value
$k_1[\frac{L^2}{mol^2 s}]$	8.7096E+08	8.94
$k_2[\frac{L^{0.4}}{mol^{0.4} s}]$	1.4791E+03	3.17
$r_{max3}[\frac{1}{s}]$	8.0E-11	-10.1
$K_{half3}^{\mathrm{C_{ORG}}}[M]$	1.2E-06	-5.9
$K_{half}^{\mathrm{NO_2Dcf}}[M]$	7.9E-10	-9.1
$K_{inhib2}[M]$	5.0E-06	-5.3
$k_3[\frac{L}{mol \cdot s}]$	2.9854e+05	5.47

Table 4: Estimated model parameters: k_1, k_2, and k_3 values correspond to MAP estimates, whereas the remaining model parameters are estimated through the corresponding empirical posterior mean value.



(a) (b)

Figure 5: Normalized concentrations of the main species involved along (a) the Nitrogen reduction cycle and (b) the Diclofenac reversible transformation pathway. Results are obtained employing the calibrated model, where the most influential parameters (k_1, k_2, and k_3) are set to corresponding MAP values, whereas the remaining ones (r_{max3}, $K_{half3}^{\mathrm{C_{ORG}}}$, $K_{half}^{\mathrm{NO_2Dcf}}$, and K_{inhib2}) are set to their posterior means. Available measurements are highlighted with diamonds. Percentage differences between Diclofenac observations (C^*) and modeling results (C) are also included.

4. Conclusions

Diclofenac (Dcf) is often detected in water resources and groundwater bodies and is increasingly recognized as a threat for the delicate balance of aquatic ecosystems, especially in view of its bioactive nature and recalcitrance. Our study provides a model and the associated operational workflow aimed at assisting the assessment of the fate of Diclofenac under uncertainty at relevant environmental conditions. The model rests on a detailed description of the molecular mechanisms associated with the biotransformation pathway of Diclofenac and is framed within a stochastic context that enables obtaining a probability distributions of target quantities conditional to available observations. We demonstrate our approach through a set of available laboratory-scale data (Barbieri et al. (2012)). Our work leads to the following major conclusions.

- Our process-based biochemical model enables us to interpret the reversible pathway exhibited by Dcf in the context of the analyzed laboratory ex-

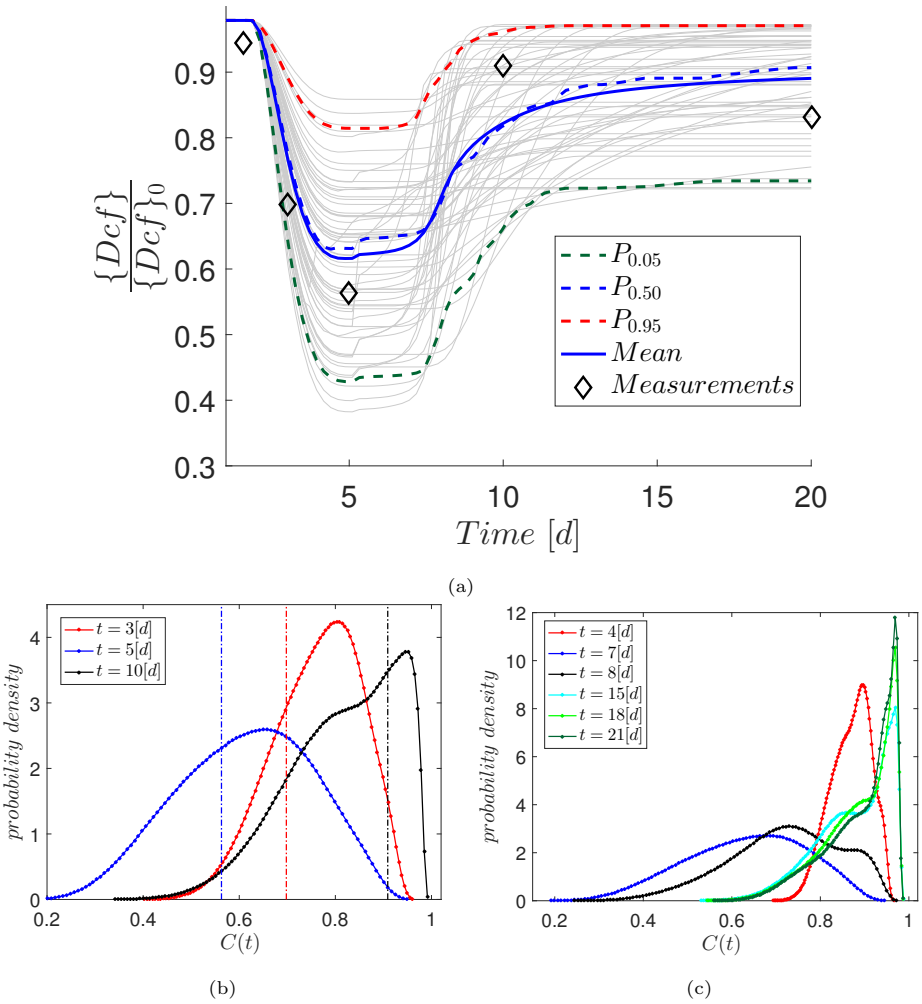


Figure 6: (a) Temporal evolution of Diclofenac concentrations obtained through a collection of exemplary model realizations associated with the posterior probability densities of model parameters (solid grey curves) together with the corresponding mean (solid blue), percentiles $P_{0.05}$, dashed green, $P_{0.5}$, dashed blue, and $P_{0.95}$, dashed red, and experimental observations (diamonds); (b) Diclofenac concentration pdfs at selected observation times together with the corresponding sampled values (vertical dashed lines); (c) Diclofenac concentration pdfs at selected unsampled times.

periments. Our conceptual model considers three subsequent phases of Dcf biotransformation, ultimately leading to the recovery of the parent compound in dissolved phase. The model couples Dcf biotransformation to the nitrogen cycle. The documented ineffectiveness of Dcf degradation under the experimental conditions investigated in this study possibly suggests the opportunity to explore diverse biodegradation pathways in future research, such as, e.g., settings associated with stronger oxidizers, eventually leading to complete mineralization of the original molecule.

- We describe selected model parameters by way of their (posterior, i.e., conditional on available data) probability distribution upon relying on an acceptance-rejection sampling algorithm. An optimal parameter combination is identified through the ensuing Maximum A Posteriori (MAP) estimates of model parameters. Relying on MAP parameter estimates yields a good agreement between model results and observations, relative residuals associated with Dcf concentrations being always smaller than 15%. Our model is seen to underestimate organic Carbon concentration, this being likely due to the fact that biomass dynamics are here neglected.

- Our results suggest that experimental observations of the kind considered here might not be exhaustive to yield sharp estimates of all of the biochemical parameters potentially affecting Dcf biotransformation. Posterior distributions associated with kinetic rates exhibit a unique peak, suggesting that optimal parameter values could be identified (in a stochastic inverse modeling context). Otherwise, the available data do not lead to a reduction of the uncertainty related to half-saturation and inhibition constants, their prior and posterior distributions being not too dissimilar. This result is likely linked to the low concentrations exhibited by the Dcf in the considered experiment, and implies that the available data do not enable one to constrain the probability distribution of inhibition and half-saturation constants associated with the *Michaelis-Menten-Monod* rates. Our analysis provides an example of the usefulness of a probabilistic framework to

identify residual model parameter uncertainty following conditioning on observations.

- Probability distributions of Dcf concentrations are obtained by propagating the obtained posterior *pdf*s of input parameters through the considered model. At sufficiently long times the *pdf*s associated with Dcf concentration values are left-skewed with a peak corresponding to a 90-100% recovery of the initial concentration. With reference to the Dcf concentrations increase, our results suggest that the back-transformation step is characterized by a fast kinetic behavior, leading to sharp variations of the Dcf with time.

Acknowledgements

This work was funded by *Bracco Imaging* (Italy).

Appendix A. Chemical structure of parent compound and transformation products

In this Section, we provide structural formulas for the main chemical species involved in the reactive processes of Diclofenac reversible pathway.

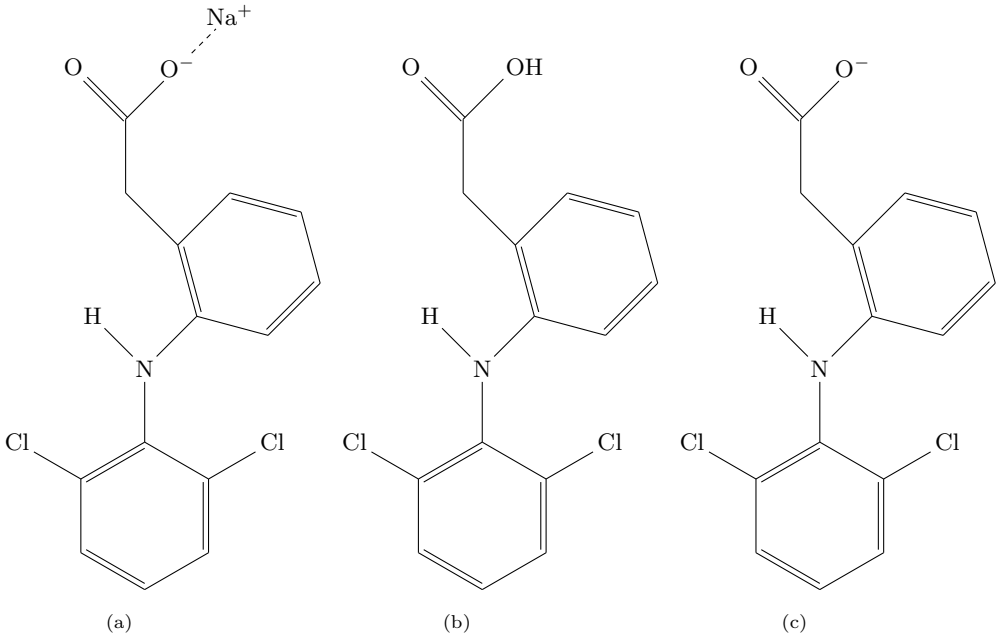


Figure A.1: Parent compound and products of its dissolution and aqueous speciation: (a) Sodium Diclofenac (NaDcf), (b) Diclofenac acid (HDcf) and (c) Diclofenac anion (Dcf^{-}).

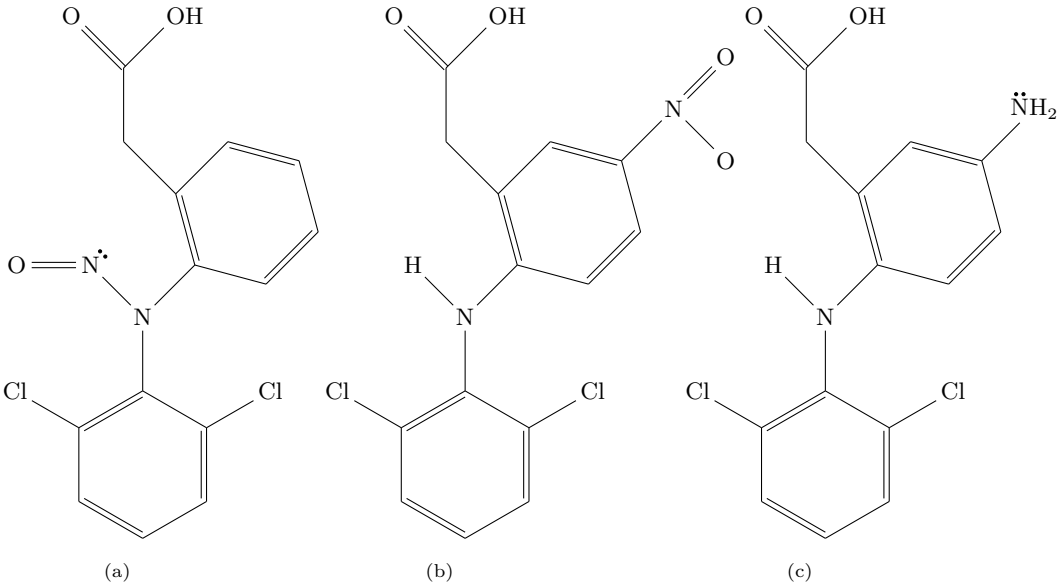


Figure A.2: Main transformation products of Diclofenac: (a) N-Nitroso-Dcf, (b) 5-Nitro-Dcf (para isomer) and (c) 5-Aminyl-Dcf (para isomer).

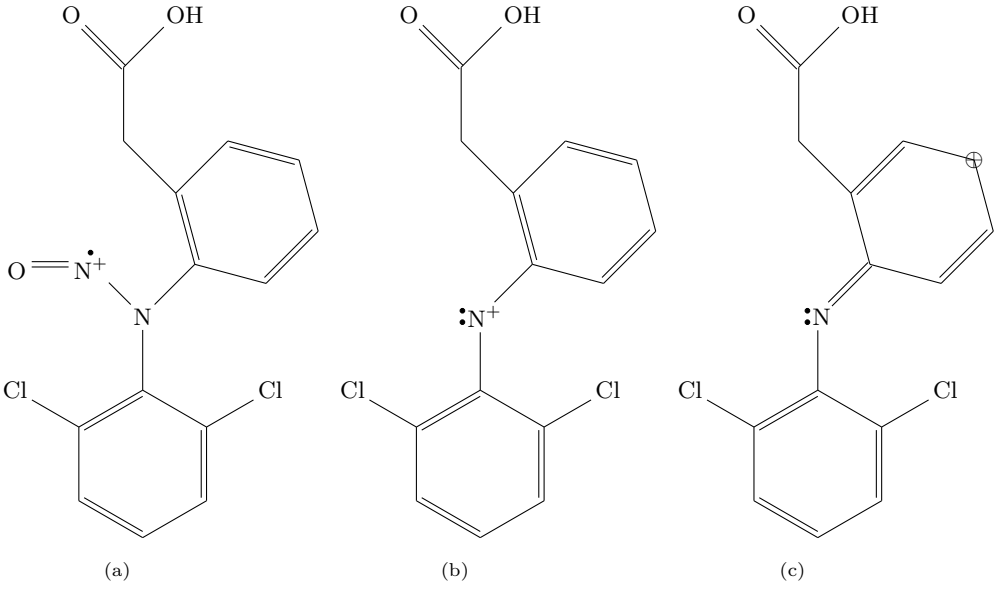


Figure A.3: Intermediates appearing in Phase (*II*) of the pathway: (a) Nitroso radical cation ($NODcf^{+\bullet}$), (b) unstable canonical form of resonance (nitrenium cation) of Dcf^{+} and (c) main canonical form of resonance (carbocation) of Dcf^{+}.

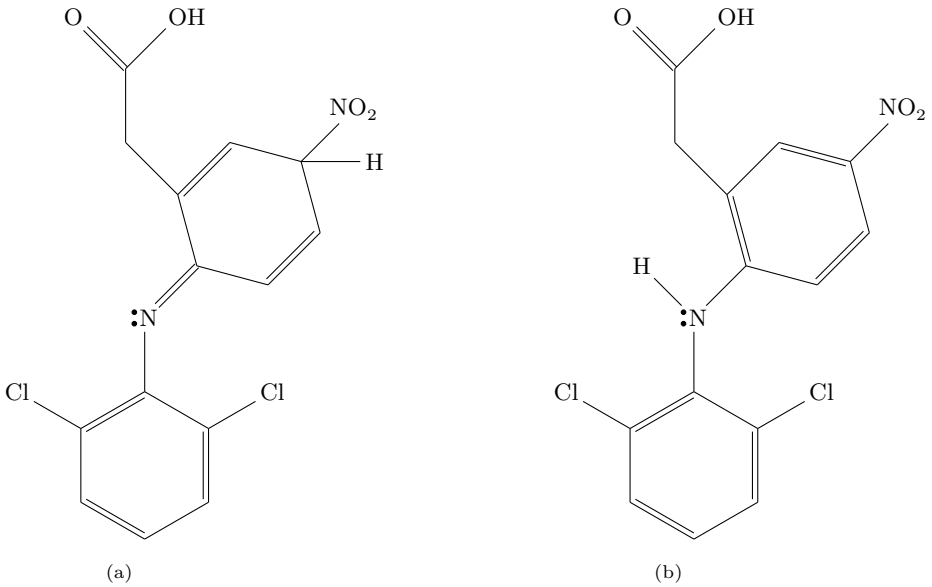


Figure A.4: Intermediates appearing in Phase (*II*) of the pathway: (a) the second structural isomer of the nitro-derivative (NO_2Dcf_{ISO2}) and (b) 5-Nitro-Diclofenac (NO_2Dcf).

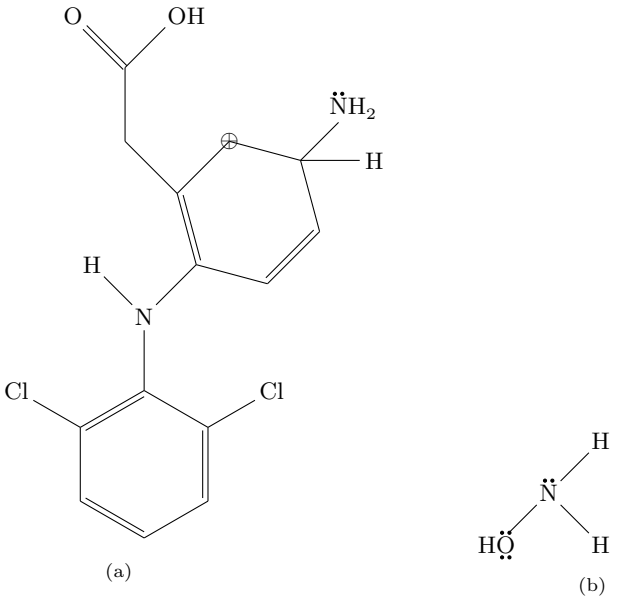


Figure A.5: Chemical species appearing during back-transformation reactions: (a) intermediate carbocation (NH_3Dcf^+) and (b) Hydroxylamine (NH_2OH).

References

Appelo, C.A.J., Postma, D., 2004. Geochemistry, groundwater and pollution. CRC press.

Barbieri, M., Carrera, J., Sanchez-Vila, X., Ayora, C., Cama, J., Köck-Schulmeyer, M., de Alda, M.L., Barceló, D., Brunet, J.T., García, M.H., 2011. Microcosm experiments to control anaerobic redox conditions when studying the fate of organic micropollutants in aquifer material. Journal of Contaminant Hydrology 126, 330–345.

Barbieri, M. Carrera, J., Ayora, C., Sanchez-Vila, X., Licha, T., Nödler, K., Osorio, V., Pérez, S., Köck-Schulmeyer, M., de Alda, M.L., et al., 2012. Formation of diclofenac and sulfamethoxazole reversible transformation products in aquifer material under denitrifying conditions: batch experiments. Science of the Total Environment 426, 256–263.

Bolstad, W.M., Curran, J.M., 2016. Introduction to Bayesian statistics. John Wiley & Sons.

Committee for Medicinal Products for Human Use (CHMP), E.M.A., 2006. Guideline on the environmental risk assessment of medicinal products for human use. EMEA/CHMP/SWP/4447/00 corr 1.

Chiron, S., Duwig, C., 2016. Biotic nitrosation of diclofenac in a soil aquifer system (katari watershed, bolivia). Science of the Total Environment 565, 473–480.

de Voogt, P., Sacher, F., Janex-Habibi, M.L., Puijker, L., Mons, M., 2007. Development of an international priority list of pharmaceuticals relevant for the water cycle. Water science and technology: a journal of the International Association on Water Pollution Research 59(1), 39–46.

Fonger, G.C., Hakkinen, P., Jordan, S., Publicker, S., 2014. The national library of medicine's (nlm) hazardous substances data bank (hsdb): background, recent enhancements and future plans. Toxicology 325, 209–216.

627 Greskowiak, J., Hamann, E., Burke, V., Massmann, G., 2017. The uncertainty
628 of biodegradation rate constants of emerging organic compounds in soil and
629 groundwater – a compilation of literature values for 82 substances. Water
630 Research 126, 122–133.

631 Kulkarni, M., Chaudhari, A., 2007. Microbial remediation of nitro-aromatic
632 compounds: an overview. Journal of Environmental Management 85, 496–
633 512.

634 Lonappan, L., Brar, S.K., Das, R.K., Verma, M., Surampalli, R.Y., 2016. Di-
635 clofenac and its transformation products: environmental occurrence and tox-
636 icity – a review. Environment International 96, 127–138.

637 Mirvish, S.S., 1975. Formation of n-nitroso compounds: chemistry, kinetics, and
638 in vivo occurrence. Toxicology and applied pharmacology 31, 325–351.

639 Nham, H.T.T., Greskowiak, J., Nödler, K., Rahman, M.A., Spachos, T., Ruste-
640 berg, B., Massmann, G., Sauter, M., Licha, T., 2015. Modeling the transport
641 behavior of 16 emerging organic contaminants during soil aquifer treatment.
642 Science of the Total Environment 514, 450–458.

643 Parkhurst, D.L., Appelo, C., 2013. Description of input and examples for
644 PHREEQC version 3: a computer program for speciation, batch-reaction,
645 one-dimensional transport, and inverse geochemical calculations. Technical
646 Report. US Geological Survey.

647 Porta, G., la Cecilia, D., Guadagnini, A., Maggi, F., 2018. Implications of
648 uncertain bioreactive parameters on a complex reaction network of atrazine
649 biodegradation in soil. Advances in Water Resources 121, 263–276.

650 Razo-Flores, E., Donlon, B., Lettinga, G., Field, J.A., 1997. Biotransforma-
651 tion and biodegradation of n-substituted aromatics in methanogenic granular
652 sludge. FEMS microbiology reviews 20, 525–538.

Rodríguez-Escales, P., Sanchez-Vila, X., 2016. Fate of sulfamethoxazole in groundwater: Conceptualizing and modeling metabolite formation under different redox conditions. Water Research 105, 540–550.

Schaffer, M., Kröger, K.F., Nödler, K., Ayora, C., Carrera, J., Hernández, M., Licha, T., 2015. Influence of a compost layer on the attenuation of 28 selected organic micropollutants under realistic soil aquifer treatment conditions: Insights from a large scale column experiment. Water Research 74, 110–121.

Small, R., 1989. Diclofenac sodium. Clinical pharmacy 8, 545–558.

Smith, M.B., 2020. March's advanced organic chemistry: reactions, mechanisms, and structure. John Wiley & Sons.

Sobol, I.M., 1998. On quasi-monte carlo integrations. Mathematics and computers in simulation 47, 103–112.

Stumm, W., Morgan, J.J., 2012. Aquatic chemistry: chemical equilibria and rates in natural waters. volume 126. John Wiley & Sons.

Wishart, D.S., Knox, C., Guo, A.C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., Woolsey, J., 2006. Drugbank: a comprehensive resource for in silico drug discovery and exploration. Nucleic acids research 34, D668–D672.

List of Supplementary Materials

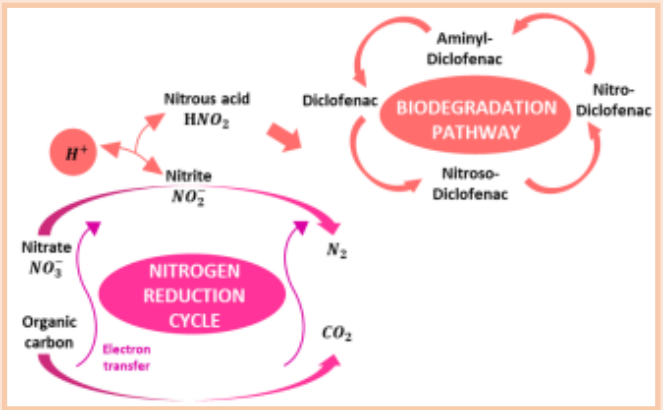
Electronic Supplementary Material (for online publication only)

S1_Detailedhydrogeochemistry_phreeqcimplementation.tex

Graphical Abstract

Formulation and probabilistic assessment of reversible biodegradation pathway of Diclofenac in groundwater

Geochemical model

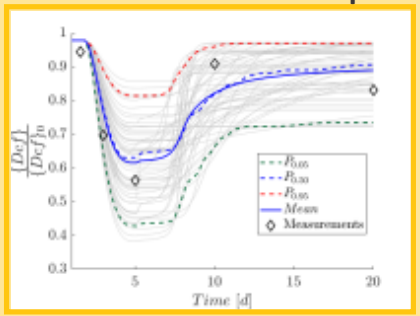


Results in a stochastic Bayesian framework

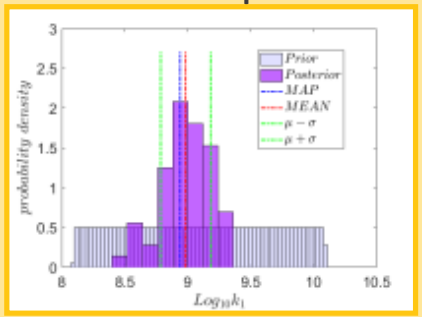
Diclofenac concentrations $= f\left(k_1, k_2, r_{max3}, K_{half3}^{C_{ORG}}, K_{half}^{NO_2Dcf}, K_{inhib2}, k_3\right)$

Outputs | Model | Parameters

Percentiles of model outputs



Estimated parameters and associated uncertainty



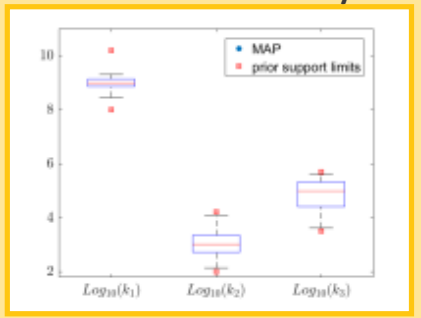


Figure 1

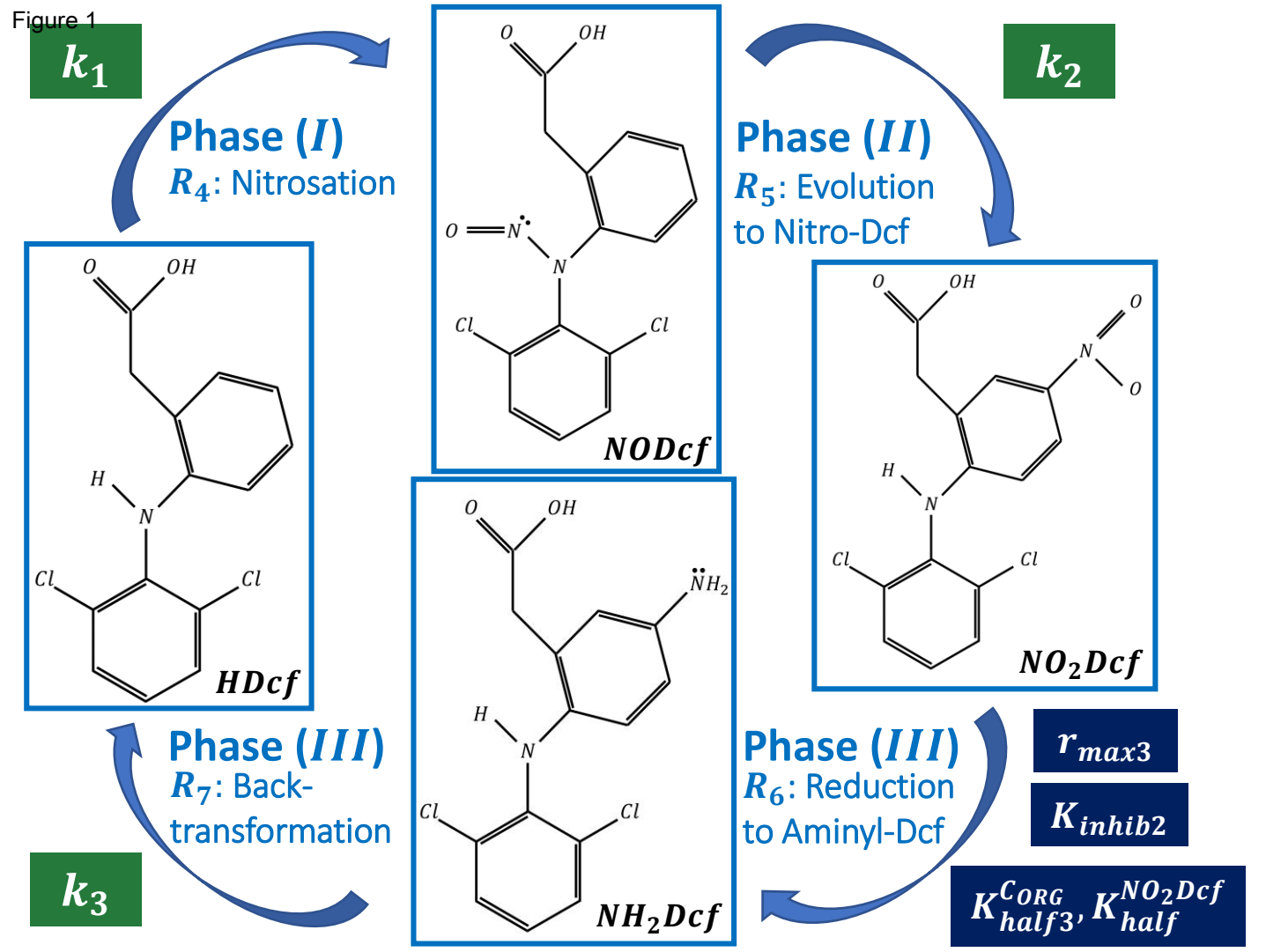
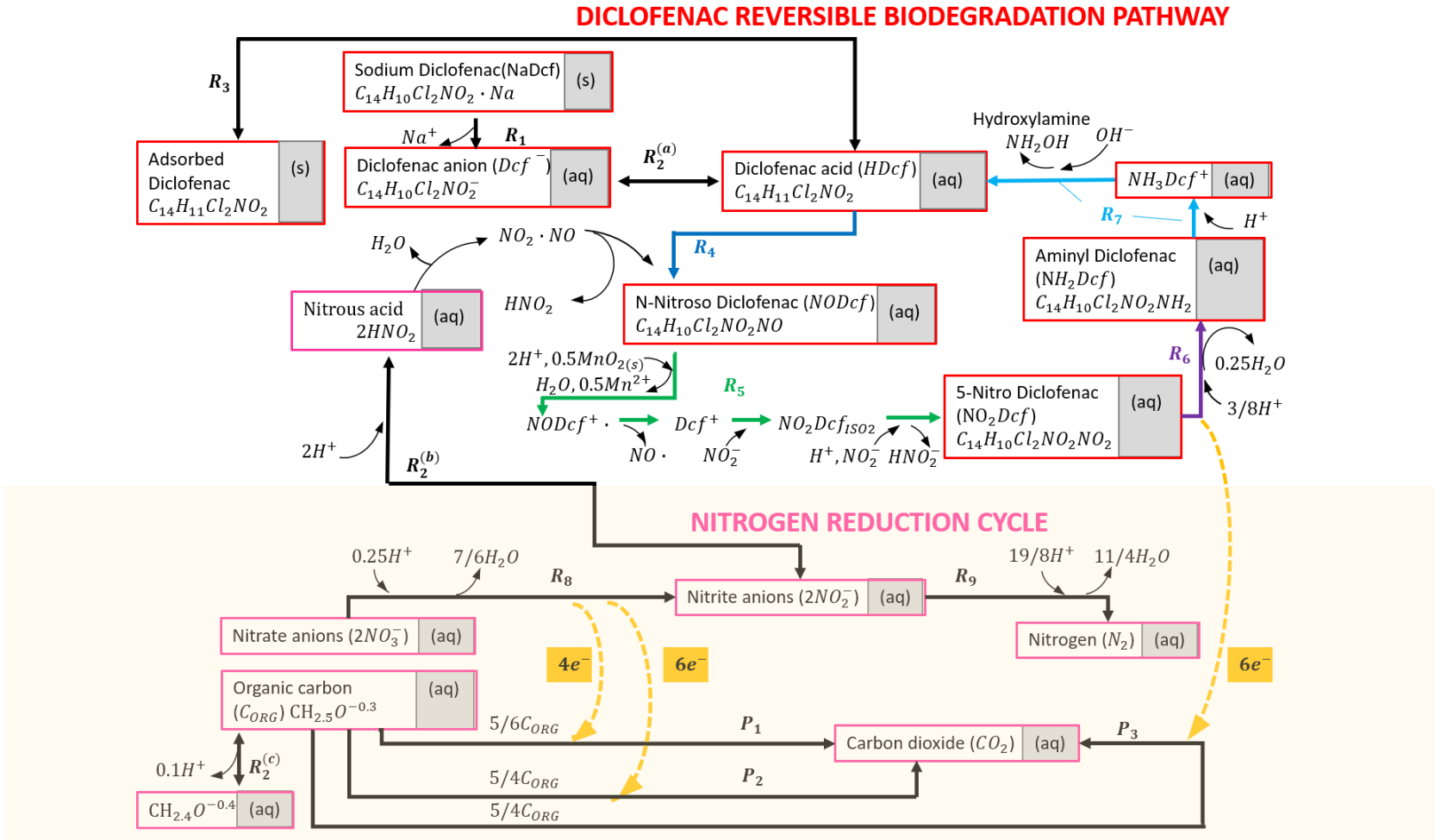

k_1
Phase (I)
R_4: Nitrosation
HDcf
NODcf
Phase (II)
R_5: Evolution to Nitro-Dcf
k_2
NO_2Dcf
Phase (III)
R_6: Reduction to Aminyl-Dcf
NH_2Dcf
Phase (III)
R_7: Back-transformation
k_3
r_{max3}
K_{inhib2}
$K_{half3}^{C_{ORG}}$, $K_{half}^{NO_2Dcf}$

Figure 2



Equilibrium driven reactions:	*Kinetic reactions:*	*Kinetic biotransformation reactions (Michaelis-Menten-Monod :*
$\boldsymbol{R_1}$: Diclofenac precipitation / dissolution $\boldsymbol{R_2^{(i)}}$: Acid dissociation / aqueous complexation of *i-th* compound $\boldsymbol{R_3}$: Diclofenac adsorption due to surface complexation	$\boldsymbol{R_4}$: Diclofenac N-nitrosation by co-metabolism $\boldsymbol{R_5}$: Nitroso to Nitro-Diclofenac evolution $\boldsymbol{R_7}$: Electrophilic Aromatic Substitution	$\boldsymbol{R_6}$: Nitro to Aminyl-Diclofenac reduction by direct metabolism $\boldsymbol{R_8}$: Reduction N(V) to N(III) $\boldsymbol{R_9}$: Reduction N(III) to N(0) $\boldsymbol{P_1}$: Organic carbon C(-0.8) oxidation to C(IV): first path $\boldsymbol{P_2}$: Organic carbon C(-0.8) oxidation to C(IV): second path $\boldsymbol{P_3}$: Organic carbon C(-0.8) oxidation to C(IV): third path