Check for updates

AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2022WR033183

On Multi-Model Assessment of Complex Degradation Paths: The Fate of Diclofenac and Its Transformation Products

Laura Ceresa[1], Alberto Guadagnini[1], Paula Rodríguez-Escales[2], Monica Riva[1], Xavier Sanchez-Vila[2], and Giovanni M. Porta[1]

[1]Department of Civil and Environmental Engineering, Politecnico di Milano, Milano, Italy, [2]Department of Civil and Environmental Engineering, Universitat Politecnica de Catalunya, Barcelona, Spain

Key Points:

- We propose a methodology to address complex multi-step transformations of diclofenac in a soil-water system
- A Multi-Model Global Sensitivity Analysis allows ranking the influence of system processes and their parameterization on model outputs
- A trade-off between model complexity and parametric uncertainty is achieved to assist stochastic model calibration

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

L. Ceresa and G. M. Porta,
laura.ceresa@polimi.it;
giovanni.porta@polimi.it

Citation:

Ceresa, L., Guadagnini, A., Rodríguez-Escales, P., Riva, M., Sanchez-Vila, X., & Porta, G. M. (2023). On Multi-Model assessment of complex degradation paths: The fate of diclofenac and its transformation products. *Water Resources Research*, *59*, e2022WR033183. https://doi.org/10.1029/2022WR033183

Received 4 JUL 2022
Accepted 28 DEC 2022

Author Contributions:

Conceptualization: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta
Data curation: Laura Ceresa
Formal analysis: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta
Funding acquisition: Alberto Guadagnini
Investigation: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta
Methodology: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta

© 2023. American Geophysical Union. All Rights Reserved.

Abstract We present a methodology to quantify the impact of model structure and parametric uncertainty on formulations targeting biotransformation processes of Emerging Contaminants in subsurface water resources. The study is motivated by recognizing that modeling of bio-mediated reactions of recalcitrant compounds in soil and aquifers is plagued by uncertainty. At the same time, process-based models often require the parameterization of complex physico-chemical processes, a situation which is exacerbated by the paucity of direct observations. Thus, assessment and formulation of modeling tools capable to balance complexity and reliability is a key challenge. The modeling strategy proposed here aims at pairing and applying a suite of quantitative tools starting from a prior diagnosis of multiple uncertainty sources and leading to parameter estimation and model selection in the presence of a limited number of observations. The methodology is illustrated through application to a multi-step, reactive scenario involving biotransformation of the pharmaceutical diclofenac (DCF) in groundwater. Our framework includes four plausible models. These are obtained through successive simplifications of a recently developed highly complex model. Such simplifications are accomplished consistent with the results of a comprehensive Multi-Model Global Sensitivity Analysis. The latter allows ranking the levels of influence of system processes on model outputs by incorporating the effects of model formulation and parametric uncertainties. The kinetic of the loop-initiating process (DCF nitrosation, linked to the temporal evolution of N-cycle components) is documented as dominating in explaining the variability of model outputs of environmental interest. Model discrimination criteria suggest that a simplified counterpart of the reference model is favored to interpret available data. Our modeling approach can assist interpretation and prototyping of a wide range of contaminant biotransformation models. The latter is a key objective also for the purpose of developing credible (environmental) risk assessment tools and designing experimental sampling campaigns.

1. Introduction

Reactive transport settings in groundwater systems involve a suite of often complex geochemical reactions. A remarkable exemplary scenario is given by the degradation of Contaminants of Emerging Concern (CECs). In this context, the evolution of one compound might result in the occurrence of numerous species. These range from the parent molecule to many metabolites, some of which being still scarcely addressed in the literature. Most of these compounds are suspected to be responsible of chronic effects on natural ecosystems and human health (Fent et al., 2006; Im et al., 2020; Kumar et al., 2010; La Farre et al., 2008). Thus, proper understanding and quantification of the fate of these types of pollutants (and transformation products) is critical from an environmental and (eco)toxicological perspective.

Reliable assessment of the fate of CECs in groundwater starts from identifying the different molecules that might arise from an individual parent. One then needs to enumerate all degradation pathways and potential processes that might lead to the presence of transformation products and by-products. The bio-geochemical conditions that might yield such degradation pathways should then be identified. A critical step of the analysis relies on (a) the selection of appropriate mathematical formulations of the processes involved in the system evolution and (b) the estimation of the values of the parameters embedded in such processes (such as, e.g., maximum degradation rates, yield coefficients or inhibition parameters; see, e.g., Rodríguez-Escales and Sanchez-Vila, 2016). Furthermore, it can be noted that the fate of CECs in groundwater generally depends on additional physico-chemical parameters, including, for example, lithologic (Reberski et al., 2022), textural and hydraulic characteristics of the soil (Farhat

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Project Administration: Alberto Guadagnini
Resources: Alberto Guadagnini
Software: Laura Ceresa
Supervision: Alberto Guadagnini, Monica Riva, Giovanni M. Porta
Validation: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta
Visualization: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta
Writing – original draft: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta
Writing – review & editing: Laura Ceresa, Alberto Guadagnini, Paula Rodríguez-Escales, Monica Riva, Xavier Sanchez-Vila, Giovanni M. Porta

et al., 2022), temperature (Greskowiak et al., 2006), redox state (Rodríguez-Escales et al., 2017), or characteristics of the microbial environment (Canelles et al., 2021).

A direct consequence of the amount of processes and parameters that need to be included in the interpretation of laboratory- and site-based experiments is that mathematical models often tend to become over-parameterized (i.e., they rely on an excessive number of parameters which may lead to data over-fitting). The situation is exacerbated by the observation that only a limited amount of (concentration) data of chemical species involved in these experimental settings is typically available. Reliable model calibration and validation are then challenging. In this context, the selection of appropriate mathematical formulations to model the considered processes is affected by uncertainty. One often wonders if introducing a large number of terms (each eventually involving a set of uncertain parameters) in a model formulation might be justified by the need of capturing as many process features as possible. Otherwise, relying on simplified approaches (with a reduced number of parameters) can sometimes still be a valuable option to explain most of the observed data. Indeed, reducing the number of uncertain model parameters (or conceptualizing them as lumped ones) would most often reduce estimation uncertainty as compared against the case of more convoluted models. In this sense, Global Sensitivity Analysis (GSA) can effectively assist to obtain a balanced trade-off between the complexity of processes description and the level of the associated parametric uncertainty. GSA can be also used to enhance our knowledge on model functioning. The latter is often a critical issue in the presence of high-dimensionality parameter spaces, as in the case of CECs degradation. The general framework typically considers a single conceptual model as explicative of the system behavior. Here, GSA allows ranking the level of influence of input parameters (or associated processes) on model outputs through the evaluation of given sensitivity indices/metrics (Dell'Oca et al., 2017; Sobol, 1993) (see, e.g., Ceriotti et al. (2018) and Elgendy and Porta (2021) for some recent applications). However, a Single-Model GSA (SM-GSA) does not allow incorporating the effects of model formulation uncertainty. This limitation is critical considering that the role of model uncertainty is increasingly recognized as key in the context of many hydrogeological and bio-geochemical systems, where several plausible models can be considered as potential candidates to interpret a given system (Chen & Ma, 2006; Hauck et al., 2008; Walker et al., 2015; X. Zhang et al., 2014). For this reason, the formulation of Single-Model sensitivity indices has been recently extended to a Multi-Model framework. The latter encompasses several plausible conceptualizations of the considered system (see, e.g., the variance-based model process sensitivity index proposed by Dai et al. (2017) and the Multi-Model *AMAM* indices introduced by Dell'Oca et al. (2020)).

In the broad context of modeling CECs degradation in groundwater, the present study is specifically focused on the non-steroidal anti-inflammatory drug diclofenac (DCF). The latter has been widely detected in various types of natural compartments, including groundwater (Jurado et al., 2019; Rozman et al., 2015; Schimmelpfennig et al., 2016). DCF can reach the groundwater compartment through various contamination routes including, for example, percolation from landfills and wastewater effluents from hospitals, nursing homes, and/or private households (Lonappan et al., 2016). Despite DCF detection is usually limited to trace concentrations (i.e., in the range between ng/L to fractions of μg/L), its potential adverse effects upon chronic exposure have been increasingly recognized (Bouly et al., 2022; Lonappan et al., 2016; Y. Zhang et al., 2021; references therein). Predicting the fate of DCF in the environment and in groundwater bodies is remarkably challenging. Most of the available experimental works report its relative concentrations with respect to input values and conjecture its fate to be mainly controlled by reversible sorption (e.g., Kiecak et al., 2019; Scheytt et al., 2006) or first-order irreversible degradation (e.g., Heberer & Adam, 2004; Schimmelpfennig et al., 2016). Nevertheless, other studies (e.g., Chiron & Duwig, 2016) evidence that DCF degradation could be described upon considering different molecular mechanisms of reaction that lead to the formation of various transformation products. Furthermore, the presence of DCF metabolites in groundwater is typically not monitored, with the exception of very few laboratory-scale experiments (e.g., Barbieri et al., 2012; Nödler et al., 2012). Otherwise, different metabolites have been detected in wastewater treatment plants during denitrification (Osorio et al., 2016). Following these findings, Chiron and Duwig (2016) postulated a degradation pathway of DCF incorporating several nitrogen-derivatives metabolites. Among these, Nitro-DCF (NO_2Dcf) is documented to be even more toxic for the environment than its parent compound, while the synergistic effects of all of the metabolites of DCF are recommended to be carefully assessed to preserve the integrity of the aquatic ecosystem (Osorio et al., 2016). Additionally, Barbieri et al. (2012) documented the transient formation of a nitrogen-derivative metabolite from DCF, as caused by the interaction between nitrite and an aromatic amine of DCF during denitrification. More recently, Ceresa et al. (2021) proposed a succession of processes involving DCF degradation and the formation of three transformation products that could

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

explain the behavior of the observed compounds in Barbieri et al. (2012). These authors also proposed and calibrated a mathematical model to describe the full reactive loop involving all compounds. As an indirect result of their modeling effort, Ceresa et al. (2021) document some difficulties in limiting the post-calibration uncertainty of some parameter estimates. This was conjectured to be possibly related to an insufficient level of information contained in the calibration data set, and/or to a lack of sensitivity of their model formulation to some uncertain parameters at times corresponding to those where data were available.

Here, we develop a comprehensive modeling strategy which is based on the implementation of a Multi-Model GSA (MM-GSA) in a general scenario of the kind discussed above. Such an analysis is conducive to enhance our knowledge on the relative impact of uncertain processes (and their parameterization) on model(s) predictions and calibration performances. It also assists in outlining a direction along which the level of complexity of target over-parameterized models can be reduced. In this context, the mathematical model proposed by Ceresa et al. (2021) to describe DCF evolution represents a stark example of a highly parameterized formulation involving several processes. Its calibration in the presence of the few available data can be seen as exemplificative of the set of challenges stemming from the assessment of CECs degradation in groundwater. Along these lines, this work is geared toward presenting a strategy to develop, calibrate and compare several plausible models to interpret environmentally relevant contamination scenarios plagued by uncertainty and data paucity. The key objective is to outline a balanced trade-off between the complexity of processes description and the associated parametric uncertainty. This objective is a key challenge also for the purpose of developing suitable environmental risk assessment tools, whose outcomes are increasingly recognized to be largely affected by model uncertainty (Chen & Ma, 2006; Ruggeri, 2009).

Our work is organized as follows. Section 2 describes the proposed methodology. For clarification purposes, our strategy is illustrated through application to the scenario involved in the DCF experiments of Barbieri et al. (2012), that are here reinterpreted in a MM context starting from the four-reactions loop proposed by Ceresa et al. (2021). Results quantifying the relative importance of model versus parameter uncertainty and the selection of the favored model formulation are presented in Section 3. Conclusions are then illustrated in Section 4.

2. Methods

We consider a general scenario involving the degradation of a parent compound into a number of metabolites under the action of a suite of bio-mediated reactions. In this framework, we identify a number N_M of alternative models that could be used in principle (i.e., prior to model calibration against observations) to quantify system dynamics. Each of these models is associated with a given level of complexity and parameterization. Here, we denote the highest-complexity model as M_1, while models M_i ($i = 2, \ldots, N_M$) represent its counterparts obtained through various levels of simplification of M_1. As sketched in Figure 1, all alternative models are collected in a model set (identified as $\boldsymbol{M}$). The latter is mathematically defined as:

$$\boldsymbol{M} = \left\{ \begin{matrix} M_1 \\ \ldots \\ M_{N_M} \end{matrix} \right\}; M_j = U\left(P_1^{(j)}, \ldots, P_{N_P^{(j)}}^{(j)}\right), \quad j = 1, \ldots, N_M. \tag{1}$$

Here, each individual model (M_j) embeds specific conceptualizations of the physico-chemical processes (P_k, with $k = 1, \ldots, N_P^{(j)}$) that govern the system evolution in model j. Accordingly, $P_k^{(j)}$ corresponds to the mathematical formulation employed in model j to describe process P_k. Note that $N_P^{(1)}$ represents the total number of system processes involved in the model set and $N_P^{(1)} \geq N_P^{(2)} \geq \cdots \geq N_P^{(N_M)}$ in our setting.

In the following, we describe the main steps of the proposed modeling approach. Our strategy aims at assessing the effects of various levels of simplification of the model with the highest number of parameters (i.e., M_1) upon relying on the rigorous comparison of the performance (in a relative sense) of all plausible models in the set in the presence of a given set of observations. Our workflow is based on studying and ranking the levels of impact of involved physico-chemical processes (which are described through given mathematical formulations embedding the action of uncertain parameters) on model output(s) of interest. For clarification purposes, the theoretical elements at the basis of the workflow (see also Figure 1) are illustrated upon considering a scenario involving

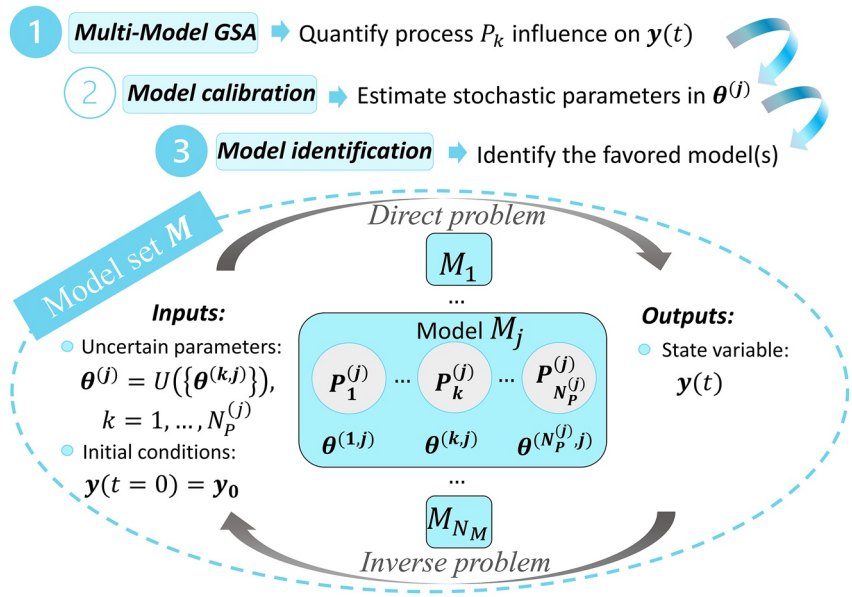

Figure 1. Main steps of the proposed model selection strategy applicable to a general model set. This is composed of multiple plausible models, each one being constructed according to a process-oriented perspective.

DCF biotransformation (which is introduced in Section 2.1). These elements are articulated through a series of steps including (I) MM-GSA, whose results can be used to assist (II) model calibration in a stochastic context, and (III) model identification. The latter allows detecting the most appropriate formulation to interpret the considered scenario under the available information content.

2.1. Characterization of Diclofenac Biotransformation in a Multi-Model Context

For the purpose of clearly enucleating all of the elements embedded in the overarching MM-GSA framework, we start by introducing the setup that will then be employed to assess our proposed method. The reference experimental data set we consider refers to the batch experiments of Barbieri et al. (2012) (their series 1 µg/L). The latter document the occurrence of DCF biotransformation under denitrifying redox conditions. A high-complexity model (labeled as M_1) is the one developed and calibrated (in a Single-Model framework) by Ceresa et al. (2021) and is taken here as exemplificative of a class of over-parameterized models. Its basic traits are illustrated in Figure 2, which embeds a summary of the involved organic compounds (DCF and three nitrogen-derivative metabolites) and inferred processes. The latter are represented by multi-stage chemical reactions involving the parent molecule (DCF acid, HDcf), its transformation products (i.e., Nitroso-DCF, NODcf; Nitro-DCF, NO_2Dcf; and Aminyl-DCF, NH_2Dcf), and several major ions dissolved in groundwater. In addition to a number of well-known reactions that are modeled by means of deterministically set parameters (i.e., equilibrium constants), this model includes four kinetic processes that are treated as stochastic (i.e., their parameters are conceptualized as random variables, consistent with the current level of knowledge about their molecular dynamics). These processes are termed P_k, $k = 1, \ldots, 4$, each occurring at a given rate of reaction (denoted as RR_k). Each rate depends on the concentrations of several solutes (i.e., the problem state variables), eventual deterministic constants and one or more uncertain parameters. The latter are listed and enclosed in green boxes in Figure 2.

The formulation of the four processes embedded in model M_1 is listed below.

- Process P_1, nitrosation of the parent compound (HDcf) into a first metabolite, NODcf:

$$\begin{aligned} &HNO_{2(aq)} + HDcf_{(aq)} \xrightarrow{RR_1} NODcf_{(aq)} + H_2O \\ &RR_1(t) = \frac{dC_{NODcf}(t)}{dt} = k_1 C_{HDcf}(t) C^2_{HNO_2}(t) \end{aligned} \quad (2)$$

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

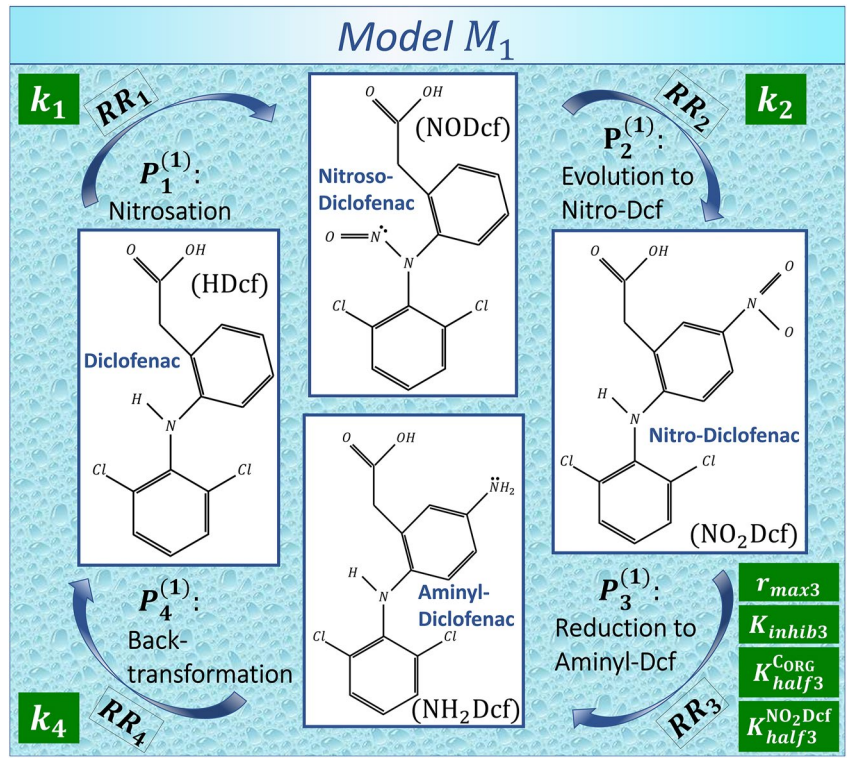

Figure 2. Geochemical model (M_1) of diclofenac bio-mediated transformations. $P_k^{(1)}$ denotes the formulation of system process k in model M_1, here consisting of complex (bio)geochemical reactions. RR_k represents the reaction rate of process k. Green boxes include the stochastic parameters as representative of specific system processes.

- Process P_2, oxidation of NODcf leading to the formation of NO_2Dcf (a second metabolite):

$$\begin{aligned} &\mathrm{NODcf_{(aq)}} + \frac{1}{2}\mathrm{MnO_{2(s)}} + 3\mathrm{H^+_{(aq)}} + 2\mathrm{NO_{2(aq)}^-} \overset{\mathrm{RR_2}}{\rightarrow} \\ &\frac{1}{2}\mathrm{Mn^{2+}_{(aq)}} + \cdot\mathrm{NO_{(aq)}} + \mathrm{H_2O} + \mathrm{NO_2Dcf_{(aq)}} + \mathrm{HNO_{2(aq)}} \\ &RR_2(t) = \frac{dC_{\mathrm{NO_2Dcf}}(t)}{dt} = k_2 \frac{C_{\mathrm{NO_2^-}}(t) C^2_{\mathrm{H^+}}(t) C_{\mathrm{NODcf}}(t)}{\sqrt{C_{\mathrm{Mn^2}} + (\mathrm{t})}} \end{aligned} \tag{3}$$

- Process P_3, reduction of the second metabolite (NO_2Dcf) to a third one (NH_2Dcf) sustained by organic matter oxidation:

$$\begin{aligned} &\mathrm{CH_4O_{(aq)}} + \frac{3}{4}\mathrm{C_2H_3O_{2(aq)}^-} + \frac{3}{4}\mathrm{H^+_{(aq)}} + 2\mathrm{NO_2Dcf_{(aq)}} \overset{\mathrm{RR_3}}{\rightarrow} \\ &\qquad \frac{5}{2}\mathrm{CO_{2(aq)}} + 2\mathrm{NH_2Dcf_{(aq)}} + \frac{1}{2}\mathrm{H_2O} \\ &RR_3(t) = \frac{1}{2}\frac{dC_{\mathrm{NH_2Dcf}}(t)}{dt} = r_{max3} \frac{C_{\mathrm{C_{ORG}}}(t)}{C_{\mathrm{C_{ORG}}}(t) + K_{half3}^{\mathrm{C_{ORG}}}} \\ &\qquad \cdot \frac{C_{\mathrm{NO_2Dcf}}(t)}{C_{\mathrm{NO_2Dcf}}(t) + K_{half3}^{\mathrm{NO_2Dcf}}} \frac{K_{inhib3}}{K_{inhib3} + C_{\mathrm{NO_2^-}}(t)} C_{\mathrm{BIOMASS}} \end{aligned} \tag{4}$$

- Process P_4, back-transformation reconverting NH_2Dcf to the parent compound (HDcf):

$$\begin{aligned} &\mathrm{NH_2Dcf_{(aq)}} + \mathrm{H_2O} \overset{\mathrm{RR_4}}{\rightarrow} \mathrm{HDcf_{(aq)}} + \mathrm{NH_2OH_{(aq)}} \\ &RR_4(t) = \frac{dC_{\mathrm{HDcf}}(t)}{dt} = k_4 C_{\mathrm{H^+}}(t) C_{\mathrm{NH_2Dcf}}(t) \end{aligned} \tag{5}$$

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Here, $C_m(t)$ denotes the aqueous concentration of species m at time t. All of these processes are driven by the presence of several major ions dissolved in groundwater. These include, for example, Mn^{2+}, H^+ and NO_2^-, and their concentrations contribute to the evolution of system reaction rates. Among these ions, nitrite (NO_2^-) and its aqueous complexation product (nitrous acid, HNO_2) play a critical role, as the overall fate of the system is ultimately governed by the underlying denitrification cycle (Ceresa et al., 2021).

In this work we consider a total of $N_M = 4$ models, that is, M_1–M_4. Each includes four processes that are modeled according to formulation $P_k^{(j)}$ (with $k = 1, \ldots, 4$, and $j = 1, \ldots, 4$). The latter are conceptualized as specific chemical reactions that are common to all candidate models. The four models feature four distinct mathematical formulations for process P_3. Otherwise, they consider the three remaining processes (i.e., P_1, P_2, and P_4) to be represented through Equations 2, 3, and 5, respectively. The rationale behind our modeling choice is based on the results of stochastic calibration of M_1 in Ceresa et al. (2021) and on the outcome of a preliminary GSA performed on the uncertain parameters of M_1 (see Supplementary Material A in Supporting Information S1). Note that the choice of employing an identical formulation for the above-described processes in all models does not constitute a limitation of our approach. The latter is fully compatible with the use of a different mathematical formulation for each of the reaction rates considered and/or with the elimination of some of the processes considered or the inclusion of new processes. Further to this, we note that our approach can be readily used when considering alternative process formulations eventually relying on mutually exclusive hypotheses on process-driving mechanisms, which might yield poor interpretation of data when eventually set in the context of a stochastic model calibration effort.

According to Equations 2, 3, and 5, each of the above-considered reaction rates involves one uncertain (random) parameter, that is, k_1, k_2, and k_4, respectively. Process P_3 is instead described through an ensemble of models characterized by decreasing complexity, that is, ranging from M_1 to M_4. This also implies a progressive reduction in the number of model parameters to be estimated. The reaction rate corresponding to process P_3 in model M_1 involves four random parameters, according to Equation 4. These terms correspond to an inhibition constant of NO_2Dcf reduction (by dissolved nitrites), K_{inhib3}; two half-saturation constants, one for the concentration of organic matter, $K_{half3}^{C_{ORG}}$, and a second one for the concentration of NO_2Dcf, $K_{half3}^{NO_2Dcf}$; and the maximum value for the reaction rate, r_{max3}. Note that biomass concentration is considered constant ($C_{BIOMASS} = 1$ mM), in agreement with Ceresa et al. (2021).

The three additional models considered together with M_1 involve successive levels of simplification to the mathematical rendering of RR_3 in Equation 4. Such simplifications are progressively applied according to the sequence described in the following (see also Supplementary Material B in Supporting Information S1):

1. The observation that the aqueous concentrations of organic carbon in Ceresa et al. (2021) are always much larger than the half-saturation constant $K_{half3}^{C_{ORG}}$ suggests that the reaction rate RR_3 could possibly be expressed through a zero-order formulation with respect to $C_{C_{ORG}}$ (see also Appelo & Postma, 2004);
2. Considering that the aqueous concentration of NO_2Dcf in Ceresa et al. (2021) is initially much smaller than the corresponding half-saturation constant ($K_{half3}^{NO_2Dcf}$) and then increases to values that are comparable (in terms of orders of magnitude) to the latter suggests the possibility to consider a first-order dependence of RR_3 versus C_{NO_2Dcf};
3. As the inhibition process of NO_2Dcf reduction by dissolved nitrites seems only relevant across a narrow temporal window (see Figure A.2 in Supplementary Material A in Supporting Information S1), it might be worthy considering a mathematical formulation that neglects such process in our model set.

The resulting expressions of RR_3 corresponding to models M_2, M_3, and M_4 read, respectively:

$$RR_3(t) = r_{max3} \frac{C_{NO_2Dcf}(t)}{C_{NO_2Dcf}(t) + K_{half3}^{NO_2Dcf}} \frac{K_{inhib3}}{K_{inhib3} + C_{NO_2^-}(t)} C_{BIOMASS}, \quad (6)$$

$$RR_3(t) = r'_{max3} C_{NO_2Dcf}(t) \frac{K_{inhib3}}{K_{inhib3} + C_{NO_2^-}(t)} C_{BIOMASS}, \quad (7)$$

$$RR_3(t) = r'_{max3} C_{NO_2Dcf}(t) C_{BIOMASS}, \quad (8)$$

$$\text{with } r'_{max3} = r_{max3} / \left(2K_{half3}^{NO_2Dcf}\right). \quad (9)$$

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Uncertain Parameters in Each Model and Process and Corresponding Intervals of Variability Which Are Considered in the Context of (I) Multi-Model Global Sensitivity Analysis and (II) Stochastic Model Calibration, and Total $N_{PAR}^{(j)}$ Number of Uncertain Parameters Encompassed in Model j

Uncertain parameters $\theta_i^{(k,j)}$							
	Model						
Process	M_1	M_2	M_3	M_4	Lower limit	Upper limit	Units
P_1	k_1	k_1	k_1	k_1	$1.2 \cdot 10^{8}$	$1.2 \cdot 10^{10}$	$\left[\frac{L^2}{mol^2 s}\right]$
P_2	k_2	k_2	k_2	k_2	$1.3 \cdot 10^{2}$	$1.3 \cdot 10^{4}$	$\left[\frac{L^{2.5}}{mol^{2.5} s}\right]$
P_3	r_{max3}	r_{max3}	-	-	$5.0 \cdot 10^{-12}$	$5.0 \cdot 10^{-10}$	$\left[\frac{1}{s}\right]$
	K_{half3}^{CORG}	-	-	-	$1.0 \cdot 10^{-7}$	$1.0 \cdot 10^{-5}$	[M]
	$K_{half3}^{NO_2Dcf}$	$K_{half3}^{NO_2Dcf}$	-	-	$7.0 \cdot 10^{-11}$	$7.0 \cdot 10^{-9}$	[M]
	K_{inhib3}	K_{inhib3}	K_{inhib3}	-	$5.0 \cdot 10^{-7}$	$5.0 \cdot 10^{-5}$	[M]
	-	-	r'_{max3}	r'_{max3}	$3.6 \cdot 10^{-3}$	$3.6 \cdot 10^{-1}$	$\left[\frac{L}{mol \cdot s}\right]$
P_4	k_4	k_4	k_4	k_4	$4.0 \cdot 10^{3}$	$4.0 \cdot 10^{5}$	$\left[\frac{L}{mol \cdot s}\right]$
Total number of uncertain model parameters							
$N_{PAR}^{(j)}$	7	6	5	4			

Note. Parameters $K_{half}^{NO_2Dcf}$, K_{inhib2} and k_3 in Ceresa et al. (2021) correspond to parameters $K_{half3}^{NO_2Dcf}$, K_{inhib3} and k_4 listed here, respectively.

The uncertain parameters that are used to describe the chemical processes in each model and the corresponding prior parameter supports (taken from Ceresa et al. (2021), see their Table 1) are listed in Table 1. Otherwise, the prior support of parameter r'_{max3}, which is introduced for the first time in the current study, is evaluated through Equation 9, which is applied to the upper and lower boundaries of the prior support of parameter r_{max3}.

The four alternative models described above are incorporated in our model set $\boldsymbol{M}$, according to:

$$\boldsymbol{M} = \begin{Bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{Bmatrix} = \begin{Bmatrix} U\left(P_1^{(1)}, P_2^{(1)}, P_3^{(1)}, P_4^{(1)}\right) \\ U\left(P_1^{(1)}, P_2^{(1)}, P_3^{(2)}, P_4^{(1)}\right) \\ U\left(P_1^{(1)}, P_2^{(1)}, P_3^{(3)}, P_4^{(1)}\right) \\ U\left(P_1^{(1)}, P_2^{(1)}, P_3^{(4)}, P_4^{(1)}\right) \end{Bmatrix}. \tag{10}$$

2.2. Step I—Process-Oriented, Multi-Model GSA

We consider N_M alternative models that can be employed to assess the behavior of an output quantity such as, for example, a problem state variable $\boldsymbol{y}$ evaluated at N_y discrete time levels (and/or locations in space). Each model operates through a generally non-linear function f_j including $N_{PAR}^{(j)}$ uncertain parameters, collected in vector $\boldsymbol{\theta}^{(j)}$, according to (see also Figure 1):

$$\boldsymbol{y} = \begin{Bmatrix} y_1 \\ \dots \\ y_{N_y} \end{Bmatrix} = f_j\left(\boldsymbol{\theta}^{(j)}\right),\ \boldsymbol{\theta}^{(j)} = \begin{Bmatrix} \theta_1^{(j)} \\ \dots \\ \theta_{N_{PAR}^{(j)}}^{(j)} \end{Bmatrix},\ j = 1, \dots, N_M, \tag{11}$$

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where $\theta_i^{(j)}$ denotes the ith random parameter of model j. As all these parameters refer to specific system processes, it is convenient to introduce an alternative notation, namely $\theta_i^{(k,j)}$, which allows to uniquely identify the kth process to which the ith parameter pertains in model j, according to:

$$\boldsymbol{\theta}^{(j)} = U\left(\boldsymbol{\theta}^{(1,j)}, \ldots, \boldsymbol{\theta}^{\left(N_P^{(j)},j\right)}\right),\ \boldsymbol{\theta}^{(k,j)} = \left\{\begin{matrix}\theta_1^{(k,j)} \\ \ldots \\ \theta_{N_{PAR}^{(k,j)}}^{(k,j)}\end{matrix}\right\},\ k = 1, \ldots, N_P^{(j)}, j = 1, \ldots, N_M. \tag{12}$$

Here, $N_{PAR}^{(k,j)}$ denotes the total number of random parameters that are included in the mathematical formulation of process k in model j.

Process-oriented MM-GSA can be performed on model set $\boldsymbol{M}$ to assess the level of influence of the involved stochastic processes on model outputs. This is done by quantifying the sensitivity of the latter toward specific system processes. Our approach rests on evaluating the first-order sensitivity indices defined in Dell'Oca et al. (2020) for system processes. Here, we refer to index *AMAE* that enables one to assess the impact of variability in process P_k on the expected value ($\mathbb{E}$) of model output Δ according to:

$$\begin{aligned} AMAE_{P_k}^{(\Delta)} = &\underbrace{\sum_{j=1}^{N_M} w(M_j)\frac{\left|\mathbb{E}\left[\Delta|M_j\right] - \mathbb{E}[\Delta]\right|}{|\mathbb{E}[\Delta]|}}_{\text{model choice contribution from } M_j:\ \Upsilon^{(j)}} \\ &+ \underbrace{\sum_{j=1}^{N_M}\left(\sum_{i=1}^{N_{PAR}^{(k,j)}} w(M_j)\frac{\mathbb{E}_{\theta_i^{(k,j)}}\left|\mathbb{E}\left[\Delta|M_j,\theta_i^{(k,j)}\right] - \mathbb{E}\left[\Delta|M_j\right]\right|}{|\mathbb{E}[\Delta]|}\right)}_{\text{overall parameter choice contribution from } M_j:\ \Xi_k^{(j)}} = \sum_{j=1}^{N_M}\left(\Upsilon^{(j)} + \Xi_k^{(j)}\right),\ k = 1, \ldots, N_P^{(1)}. \end{aligned} \tag{13}$$

Note that Equation 13 is a streamlined formulation of Equation (C2) of Dell'Oca et al. (2020) (see our Supplementary Material C in Supporting Information S1 for the full derivation). Here, Δ represents the selected quantity of interest; $N_P^{(1)}$ denotes the total number of processes involved in the model set (this corresponding to the formulation of the high-complexity model M_1, as previously discussed), and $w(M_j)$ (with $j = 1, \ldots, N_M$) are the prior weights, which are considered uniform for all competing models in our study (i.e., $w(M_j) = 1/N_M = 1/4$), consistent with the absence of any prior indication eventually favoring any particular candidate of the set. $\mathbb{E}\left[\Delta|M_j\right]$ and $\mathbb{E}[\Delta]$ represent the unconditional expectations of Δ associated with model j and with the whole ensemble of possible outcomes of all models considered, respectively, while $\mathbb{E}\left[\Delta|M_j,\theta_i^{(k,j)}\right]$ denotes expectation of Δ conditional on the ith parameter employed in model j to describe process k (i.e., $\theta_i^{(k,j)}$). Notation $\mathbb{E}_{\theta_i^{(k,j)}}$ indicates that the corresponding expectation is taken across the prior support $\Gamma_i^{(k,j)}$ of $\theta_i^{(k,j)}$. Note that here we evaluate the above mentioned statistical moments upon numerical Monte Carlo sampling of the model and parameter space. From a practical standpoint, the prior parameter space can be (numerically) sampled through, for example, a quasi-Monte Carlo scheme (Sobol, 1998). Details on the numerical procedure for evaluating the above-defined conditional expectations are available in Supplementary Material C in Supporting Information S1.

Ranking of uncertain model processes (and associated stochastic parameters) is accomplished on the basis of the relative magnitude of the corresponding *AMAE* indices evaluated according to Equation 13. In this context, the larger the value attained by $AMAE_{P_k}^{(\Delta)}$, the more sensitive $\mathbb{E}[\Delta]$ on the formulation of process P_k (including the choice of mathematical conceptualization and the associated uncertain parameter values). Accordingly, our MM-GSA allows appreciating the relative weights of model formulation and parametric uncertainties (encapsulated in terms $\Upsilon^{(j)}$ and $\Xi_k^{(j)}$ in Equation 13, respectively) in terms of their contribution to explain the variability of the expectation of model outputs. Note also that Equation 13 embeds two distinct summation terms, that is, over j and i, respectively. The former is associated with the observation that $AMAE_{P_k}^{(\Delta)}$ includes distinct contributions (each encompassing both model and parameter uncertainty) from different competing models. The latter enables one to account for the cumulative effect of the uncertainty of multiple parameters which are eventually embedded in the formulation of process k. This is, for example, the case of process P_3 in our setting (in all model candidates except for M_4), while each of the remaining processes (i.e., P_1, P_2, and P_4) depends on a unique stochastic parameter (see Table 1).

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.2.1. Quantities of Interest

In this study, we rely on two different target quantities corresponding to the quantity Δ in Equation 13. The approach is applied to the ensemble of models described in Section 2.1. All quantities are numerically evaluated by implementing the considered geochemical models in the PHREEQC environment (version 3.6.2) (Parkhurst & Appelo, 2013).

As a first quantity of interest, we consider the classical sum of the squared differences between corresponding model results (y_m) and observed data (y_m^*), with $m = 1, \ldots, N^* \leq N_y$, that is,:

$$\varphi = \sum_{m=1}^{N^*} (y_m - y_m^*)^2. \tag{14}$$

Here, N^* denotes the number of available measurements in the calibration data set. Under some assumptions, minimizing φ yields Maximum Likelihood (ML) estimates of parameter values, that is, considering φ is tantamount to minimize the Negative-Log-Likelihood criterion, $NLL^{(j)}$ (Carrera & Neuman, 1986; see also Section 2.3).

In our DCF biotransformation models, the data set comprises a collection of $N^* = 14$ (normalized) concentration values associated with several chemical species dissolved in the aqueous phase (Ceresa et al., 2021). Note that among these, only $N^{**} = 5$ measurements are associated with DCF, the remaining ones being related to other chemical species relevant to denitrification. Concentration values, hereafter denoted as $C^*_{spec,m}$ (with $m = 1, \ldots, 14$), were collected at five selected time levels ($t^* = [1.8; 3; 5; 10; 20]$ days; see Barbieri et al., 2012) and normalized in Ceresa et al. (2021) against initial concentrations of target master species. The corresponding aqueous concentrations evaluated through the alternative models are termed $C_{spec,m}$, with $m = 1, \ldots, 14$. Normalized measured and simulated concentrations form the entries of the following vectors:

$$\boldsymbol{C}_{\boldsymbol{spec}} = \begin{Bmatrix} C_{spec,1} \\ \ldots \\ C_{spec,14} \end{Bmatrix} \quad ; \quad \boldsymbol{C}^*_{\boldsymbol{spec}} = \begin{Bmatrix} C^*_{spec,1} \\ \ldots \\ C^*_{spec,14} \end{Bmatrix}. \tag{15}$$

As a second quantity of interest, we focus on the behavior of the temporal history of DCF concentration ($C(t)$), which represents the key output ($\boldsymbol{y}$) of our geochemical models. Here, we consider $N_y = 200$ discrete values of model-based DCF concentrations collected in vector $\boldsymbol{C}$:

$$\boldsymbol{C} = \begin{Bmatrix} C_1 \\ \ldots \\ C_{200} \end{Bmatrix}. \tag{16}$$

Note that MM-GSA for quantity $\boldsymbol{C}$ relies only on DCF concentrations that are evaluated through the considered geochemical models.

Focusing on these two quantities enables one to address the effects of parameter (and process) uncertainty on both (a) calibration-related performances and (b) intrinsic features of the models prior to calibration. The latter aspect is related to the diagnosis of model functioning and processes interaction, independent of data availability. We further note that the specific model outputs toward which targeting a Global Sensitivity Analysis is a choice of the modeler. Accordingly, the methods we propose can be readily tailored to fully consider problem-specific and goal-oriented requirements.

2.3. Step II—Maximum Likelihood Model Calibration

Upon completion of MM-GSA, one can employ the resulting information to assist model calibration/inversion under scarcity of data. The identification of model parameters and/or processes associated with low impact on model outputs can lead to exclude these from a stochastic model calibration. In this study (and without loss of generality) we choose to perform inverse modeling through ML (see Carrera & Neuman, 1986). To do so, we

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

rely on the PEST (Model Independent Parameters Estimation and Uncertainty Analysis) suite (Doherty, 2015a) coupled with the PHREEQC solver, the latter being employed to simulate the proposed DCF biotransformation models (Parkhurst & Appelo, 2013).

ML calibration of model j aims at minimizing $\mathrm{NLL}^{(j)}$, defined as:

$$\mathrm{NLL}^{(j)} = -2\log_{10} L\left(\boldsymbol{\theta}^{(j)} | \boldsymbol{y}^{*}\right). \tag{17}$$

Here, $L(\boldsymbol{\theta}^{(j)}|\boldsymbol{y}^{*})$ represents the likelihood of the parameters enclosed in vector $\boldsymbol{\theta}^{(j)}$ conditional to knowledge of a number N^{*} of available information collected in $\boldsymbol{y}^{*}$ (i.e., data). Under the assumptions that (a) the variance σ_j^2 associated with prior measurement errors in model j is the same for all data; (b) such errors are mutually independent, and (c) observations and uncertain parameters are distributed according to a multi-Gaussian probability density function, Equation 17 becomes:

$$\mathrm{NLL}^{(j)} = N^{*}\log_{10}(2\pi) + N^{*}\log_{10}\left(\sigma_j^2\right) + \frac{\varphi}{\sigma_j^2}. \tag{18}$$

Notice that an unbiased estimate $\hat{\sigma}_j^2$ of the variance of prior errors is represented by the ratio between the minimized sum of squared residuals in model j(hereafter termed as φ_{MIN}) and the number of available observations, according to:

$$\hat{\sigma}_j^2 = \frac{\varphi_{MIN}}{N^{*}}. \tag{19}$$

Available data correspond to $N^{*} = 14$ (normalized) concentration values forming the entries of vector $\boldsymbol{C}^{*}_{\boldsymbol{spec}}$, as defined in Equation 15.

2.4. Step III—Models Comparison and Identification

This step relies on the evaluation of posterior model probabilities resulting from the application of formal model identification criteria. Such an approach is largely employed to discriminate among several plausible models (e.g., Janetti et al., 2012; and references therein). For the purpose of our application we rely on the Kashyap Information Criterion, KIC (Kashyap, 1982), which is defined as:

$$\mathrm{KIC}^{(j)} = \mathrm{NLL}^{(j)}_{MIN} + N^{(j)}_{PAR}\ln\left(\frac{N^{*}}{2\pi}\right) - \ln\left(\det \boldsymbol{Q}^{(j)}\right). \tag{20}$$

Here, $\mathrm{NLL}^{(j)}_{MIN}$ is the minimized Negative-Log-Likelihood of model M_j upon calibration; and $\boldsymbol{Q}^{(j)}$ represents the Cramer-Rao Lower Bound of the parameter covariance matrix associated with calibrated model M_j, which is in turn related to the Fisher information matrix. Entries of the latter depend on the Jacobian matrix, whose elements correspond to the derivatives of the target state variable with respect to model parameters (see, e.g., Carrera & Neuman, 1986). This enables one to discriminate among various models on the basis of their quality of fit to data, number of parameters, and quality of the available data and parameter estimates. In this context, relying on $\det \boldsymbol{Q}^{(j)}$ enables KIC to balance parsimony with the expected information content and to favor the model that is least probable (in an average sense) of being incorrect (Janetti et al., 2012; Ye et al., 2008; Ye et al., 2010; references therein). An extensive discussion of these model selection criteria is presented in Ye et al. (2008), Ye et al. (2010), and references therein.

The KIC-based model identification criterion relies on evaluating the differences ($\delta\mathrm{KIC}^{(j)}$) between the KIC values associated with each competing model and the minimum KIC across the model set, according to:

$$\delta\mathrm{KIC}^{(j)} = \mathrm{KIC}^{(j)} - \min_{j=1,\ldots,N_M}\left(\mathrm{KIC}^{(1)},\ldots,\mathrm{KIC}^{(N_M)}\right). \tag{21}$$

The posterior probability associated with each (calibrated) model can then be evaluated as:

$$\mathbb{P}^{(j)} = \frac{\exp\left(-\frac{1}{2}\delta\mathrm{KIC}^{(j)}\right)\mathbb{P}^{(j)}_{prior}}{\sum_{j=1}^{N_M}\left[\exp\left(-\frac{1}{2}\delta\mathrm{KIC}^{(j)}\right)\mathbb{P}^{(j)}_{prior}\right]}. \tag{22}$$

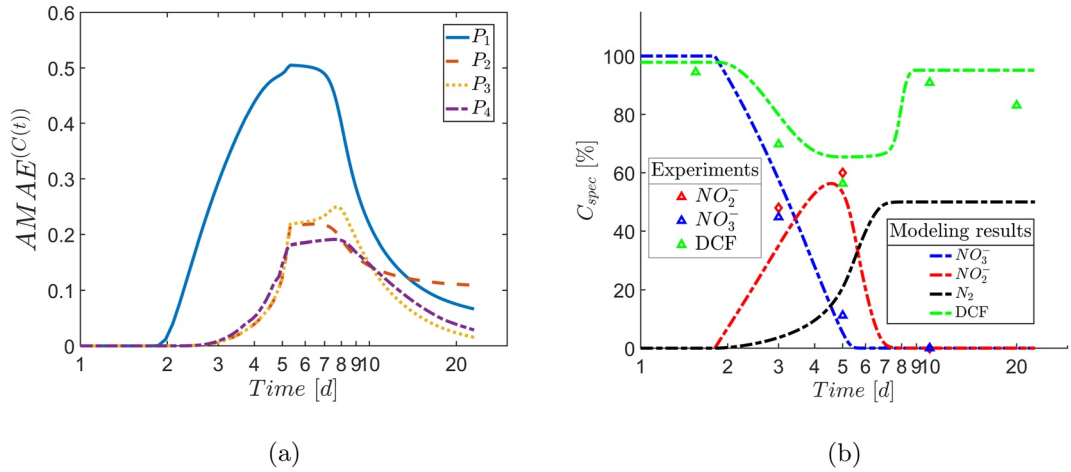

Figure 3. (a) Multi-model $AMAE_{P_k}^{(C(t))}$ sensitivity indices associated with Diclofenac (DCF) concentrations $C(t)$ and (b) temporal evolution of concentrations of DCF and nitrogen-based species that are relevant in the context of denitrification (see Ceresa et al., 2021).

Note that in our setting, all model candidates are associated with identical (uniform) prior probabilities, that is, $\mathbb{P}_{prior}^{(j)} = 1/N_M = 1/4$.

Values of posterior probabilities evaluated with KIC yield a model ranking that combines the quality of the estimates with model structure and complexity.

3. Results and Discussion

3.1. Process-Oriented, Multi-Model GSA

3.1.1. GSA of Diclofenac Concentration Values

We start by analyzing the temporal segments where specific processes affect the target output of the geochemical models (i.e., the DCF concentration history, $C(t)$). Figure 3a depicts the temporal patterns of the $AMAE_{P_k}^{(C(t))}$ family of curves (one curve per each stochastic process P_k, with $k = 1, \ldots, 4$), as assessed according to Equation 13. Here, the main focus is on model functioning and process interactions across the whole temporal window of the experiments of Barbieri et al. (2012). Nitrosation (i.e., P_1) is clearly evidenced as a dominant process along the whole time window. Its relevance steeply increases at early times. It then attains a nearly horizontal plateau at intermediate times and decreases at late times. The remaining processes (i.e., P_2–P_4) are less relevant while non-negligible, in agreement with the intermediate values attained by the corresponding $AMAE_{P_k}^{(C(t))}$ curves (with $k = 2, 3, 4$). To elucidate the geochemical dynamics underpinning this behavior, Figure 3b depicts the evolution of concentrations of the nitrous compounds involved in the denitrification cycle together with the DCF concentration curve obtained through calibration of model M_1 in Ceresa et al. (2021).

Concentration of nitrite (NO_2^-) is key in driving the fate of DCF in groundwater. Indeed, the concentration of nitrous acid (HNO_2, the product of equilibrium-driven aqueous complexation of NO_2^-) acts as a driving term for process P_1 (whose reaction rate is expressed through Equation 2). Accordingly, the nitrosation process is initiated when nitrates (NO_3^-) evolve to nitrites, whose concentration is depicted in Figure 3b (red curve). When the dissolved amount of NO_2^- displays a peak (at $t \approx 5$–6 days), P_1 is characterized by its highest degree of impact on $C(t)$ across the temporal domain considered. The importance of P_1 sharply decreases for $t > 8$ days. The remaining processes activate at $t \approx 2.5$ days, these being driven by concentrations of NODcf, a reaction product of P_1. Here, all the subsequent processes to P_1 (along the reactive loop) start to become more influential to the fate of DCF (see Figure 3a). Index $AMAE_{P_2}^{(C(t))}$ stabilizes to a nearly constant level at approximately 5 days. In this intermediate range of times (i.e., $t \in [3; 7]$ days), process P_3 is subject to the inhibition (in all models but M_4) by dissolved nitrites, that are characterized by a higher priority of reduction than NO_2Dcf (which is a product of P_2 and a reactant associated with P_3). When the inhibition effect on P_3 relaxes (after approximately 7−8 days, i.e., when the concentration of NO_2^- approaches zero), the nitrogen-derivatives of DCF are quickly reconverted to

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

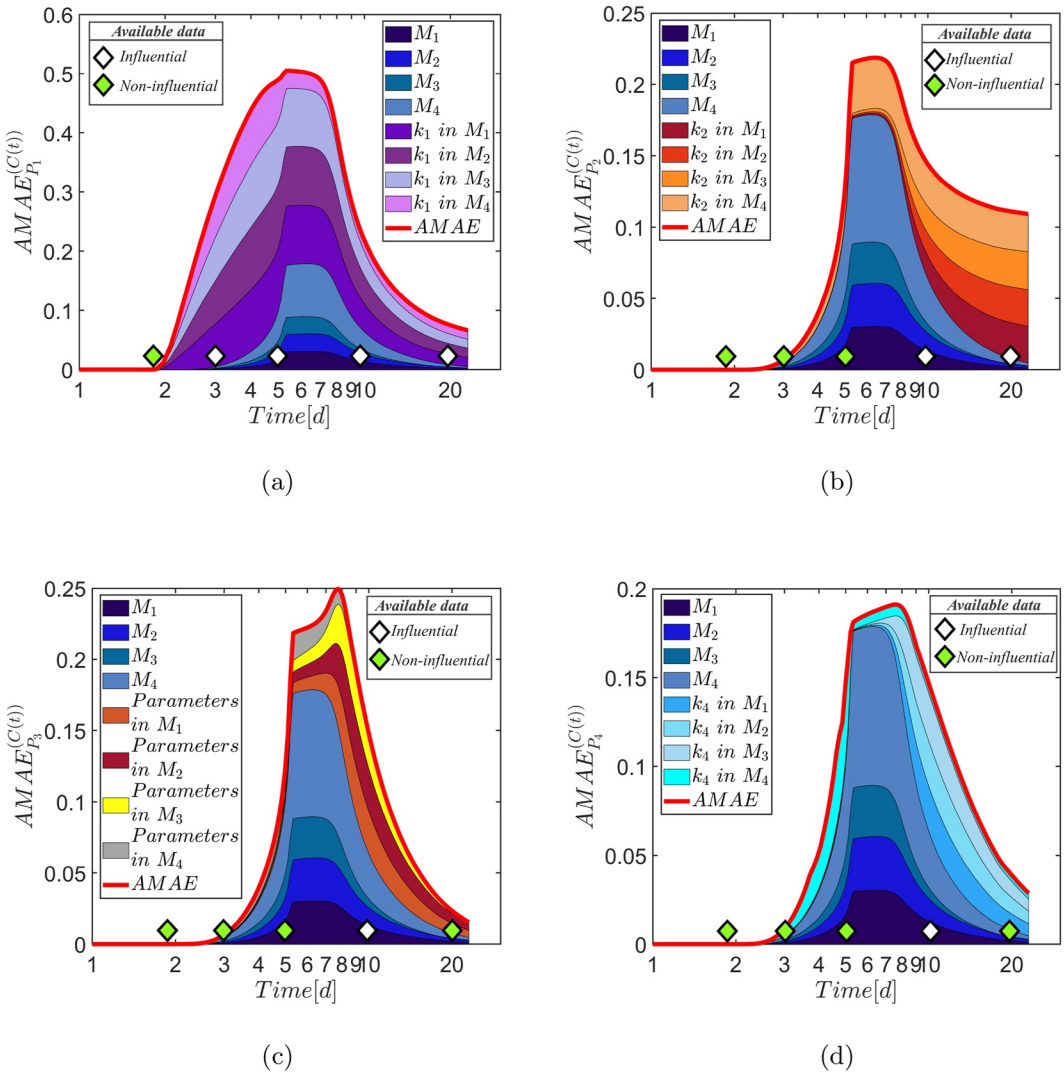

Figure 4. Multi-Model $AMAE_{P_k}^{(C(t))}$ sensitivity indices associated with $C(t)$ for (a) P_1, (b) P_2, (c) P_3, and (d) P_4. Symbols mark the temporal distribution of samples associated with the experiments of Barbieri et al. (2012); green symbols correspond to time levels where $AMAE_{P_k}^{(C(t))} < 0.05$ and/or the corresponding overall parameter choice contribution $\sum_j \Xi_k^{(j)}(t) < 0.025$ for each P_k.

the parent compound (through the simultaneous occurrence of processes P_1–P_4) and the reversible transformation cycle ends. Consequently, the importance of all processes on model output starts decreasing, as shown by the behavior of all $AMAE_{P_k}^{(C(t))}$ indices (with $k = 1, \ldots, 4$) that drop to low values for $t > 10$ days.

Figure 4 displays the distinct contributions (i.e., due to model formulation and overall parametric uncertainty within each model) to the total $AMAE_{P_k}^{(C(t))}$ values depicted in Figure 3a for $k = 1, \ldots, 4$. In this context, we also highlight the temporal locations t^* where experimental data from Barbieri et al. (2012) are available for DCF. For P_1 (see Figure 4a), the magnitude of the contributions to $AMAE_{P_1}^{(C(t))}$ ascribed to model uncertainty and to the uncertainty related to parameter k_1 is similar across the whole time window. The four models feature distinct formulations for P_3 and model contributions become relevant (for all processes) only when P_3 activates. Otherwise, for early times ($t < 3$ days) DCF concentration is sensitive only to the value assumed by parameter k_1. For the other three processes considered (i.e., P_2 to P_4; see Figures 4b–4d), sensitivity toward model formulation is larger than that associated with uncertain parameters at times comprised between 3 and 8 days. Both model and parameter choice contributions to $AMAE^{(C(t))}$ tend to drop for $t > 8$ days, consistent with the trends exhibited by $AMAE^{(C(t))}$ (i.e., red curves in Figure 4). Specifically, model choice contributions to $AMAE^{(C(t))}$ display a sharp decrease after 8 days for all processes, while sensitivity to their parameters appears slightly more persistent in

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

time. The latter feature is extremely marked for process P_2, where sensitivity to k_2 becomes notably dominant over the effect of model structure for $t > 10$ days. Consequently, $AMAE_{P_2}^{(C(t))}$ does not exhibit any sharp late-time decay, as shown in Figure 4b.

Figure 4 also allows ascertaining the results of the sensitivity analysis at the specific time levels where experimental data are available. In this context, the analysis is useful to assist identification of temporal regions where observations can provide useful information to the purpose of model calibration (i.e., to constrain the estimation uncertainty associated with the stochastic parameters). We set a (minimum) threshold value *thr* of $AMAE_{P_k}^{(C(t))}$ below which we consider process P_k as uninfluential on the model output of interest (i.e., $C(t)$). Accordingly, any observation (i.e., DCF concentration measurement $C^*(t^*)$) collected at time levels such that $AMAE_{P_k}^{(C(t^*))} < thr$ is deemed as uninfluential for the purpose of estimating the stochastic parameters of P_k through model calibration. Otherwise, data associated with time levels such that $AMAE_{P_k}^{(C(t^*))} \geq thr$ can be seen as informative, provided that the relative impact of parametric uncertainty is sufficiently significant at these times (i.e., term $\sum_j \Xi_k^{(j)}(t^*)$ in Equation 13 should at least exceed a minimum (threshold) value thr_2). For the purpose of our analysis, we set here $thr = 0.05$ and $thr_2 = 0.025$. The former roughly corresponds to 10% of the peak value for $AMAE_{P_1}^{(C(t))}$ (P_1 being the dominant process), while the latter is about 50% of *thr*. Thus, data collected outside temporal windows where a given process is influential to $C(t)$ are not expected to provide useful information to constrain the process parameterization. In this sense, Figures 4c and 4d reveal that only one of the observations included in the data set of Barbieri et al. (2012) (i.e., the one collected at $t^* = 10$ days) is enclosed in the time intervals where processes P_3 and P_4 are influential, respectively. Moreover, the impact of model structure (i.e., $\sum_j \Upsilon^{(j)}$ in Equation 13) is comparable to the one due to parametric uncertainty (i.e., $\sum_j \Xi_k^{(j)}$ in Equation 13, with $k = 3, 4$) at this specific time. These results suggest that parameters related to processes P_3 and P_4 scarcely influence the expected value of concentrations at the time levels where concentration data have been collected by Barbieri et al. (2012). Therefore, these parameters may be difficult to estimate upon relying on the available data set. Conversely, expected values of DCF concentrations display a marked sensitivity to parameters k_1 and k_2 in the initial and final phases of the experiments of Barbieri et al. (2012), respectively, as shown in Figures 4a and 4b.

The type of analysis carried out in the previous paragraphs might also assist the design of future (experimental) sampling campaigns. Indeed, our proposed MM-GSA enables one to point out the most advantageous (time) regions where experimental measurements could be collected (i.e., where (a) process-based *AMAE* indices attain large values and (b) the latter are dominated by parameter choice contributions).

3.1.2. GSA of the Objective Function Associated With Model Calibration

We analyze here the impact of model and parametric uncertainties on the average value of the objective function employed for the purpose of model calibration (i.e., φ in Equation 14). The values of $AMAE_{P_k}^{(\varphi)}$ (with $k = 1, \ldots, 4$) are equal to 0.3510, 0.1356, 0.1451, and 0.1320 for processes P_1, P_2, P_3, and P_4, respectively. Here, we note that $AMAE_{P_1}^{(\varphi)}$ is larger than $AMAE_{P_k}^{(\varphi)}$ ($k = 2, 3, 4$). This result suggests that nitrosation (i.e., process P_1) has the largest influence on the expected value of the objective function. Otherwise, the remaining processes (i.e., P_2–P_4) cannot be completely disregarded, even as they appear to be less relevant than P_1. Their effects remain still appreciable, the corresponding values of $AMAE_{P_k}^{(\varphi)}$ (with $k = 2, 3, 4$) exhibiting roughly similar values, consistent with results depicted in Figure 3a.

Individual contributions of model and parametric uncertainties can be assessed separately through the analysis of the terms appearing in Equation 13. Figure 5 displays process-specific results and highlights the contributions associated with the choice of each model formulation and involved uncertain parameters to $AMAE^{(\varphi)}$. Note that the sum of the heights of each of the vertical bars in each Figure referred to process k corresponds to the overall $AMAE_{P_k}^{(\varphi)}$ value (i.e., 0.3510, 0.1356, 0.1451, 0.1320 for $k = 1, 2, 3, 4$, respectively, as stated above).

Models M_1 and M_2 exhibit virtually indistinguishable effects of model and parameter choice contributions to $AMAE^{(\varphi)}$ for all processes, given that the contribution of parameter $K_{half3}^{\mathrm{C_{ORG}}}$ in M_1 is negligible (see also Supplementary Material A in Supporting Information S1). The remaining parameters of P_3 are involved in both models M_1 and M_2, where they display a very similar influence on φ. Otherwise, M_3 is associated with a marked reduction in its overall contribution (i.e., due to model and parametric uncertainties) to $AMAE_{P_3}^{(\varphi)}$ as compared against any of the other models (i.e., $\Upsilon^{(3)} + \Xi_3^{(3)} < \Upsilon^{(j)} + \Xi_3^{(j)}$, $j = 1, 2, 4$; see Figure 5c). The most simplified model

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

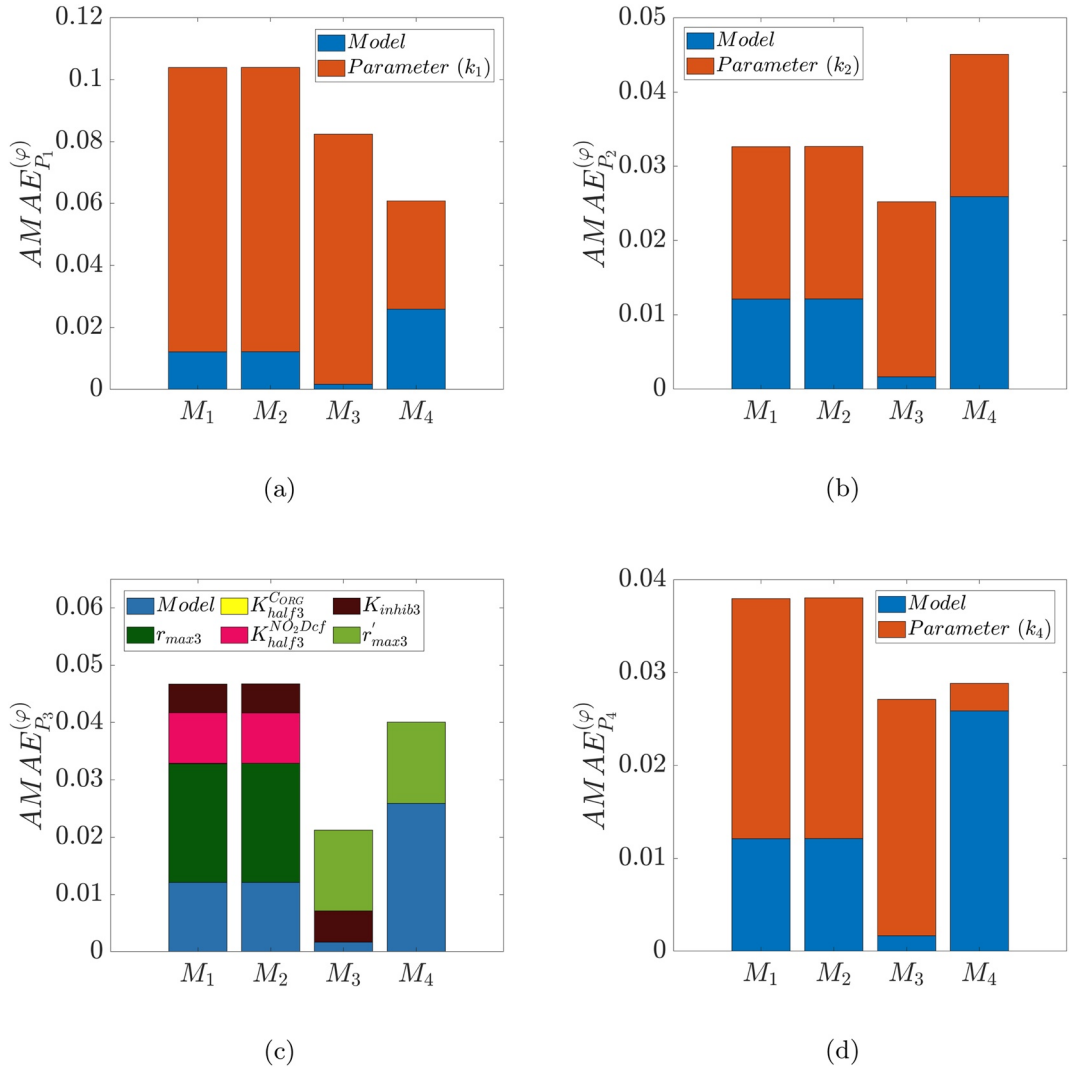

Figure 5. Values of $AMAE_{P_k}^{(\varphi)}$ corresponding to each of the models analyzed in our Multi-Model context, that is, (a) $k = 1$, (b) $k = 2$, (c) $k = 3$, and (d) $k = 4$.

candidate considered (i.e., M_4) is characterized by a notable increase of the weight associated with model formulation uncertainty as compared against that ascribed to its overall parametric uncertainty. These two results can be explained upon analyzing the impact of model structure on the distribution of the values of φ obtained through our Monte Carlo results across each model parameter space, as shown in Figure 6.Here, we observe that the range of φ values decreases from M_1 to M_4 as a result of model simplification. Model M_4 is seen to yield a much narrower distributions of φ values if compared with M_1 and M_2. Figure 6 also shows that the average value of φ for M_3 is close to the unconditional MM average $\mathbb{E}[\varphi]$. This result is at the basis of the low value of the model contribution to $AMAE_{P_k}^{(\varphi)}$ ($k = 1, \ldots, 4$) observed when considering the results of M_3 (see Figure 5).

3.2. Maximum Likelihood Model Calibration

ML model calibration is performed to estimate unknown parameters in each of the models considered. Considering the paucity of observations available in Barbieri et al. (2012), the simultaneous estimation of all stochastic parameters included in Table 1 is fraught with difficulties and does not yield acceptable results. In particular, the condition numbers associated with the parameter covariance matrix ensuing model calibration attain large values for all models (results not shown). Such behavior is likely due to the low information content that the available data set contributes toward specific system processes (i.e., P_3 and P_4; see also Section 3.1.1). Following

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

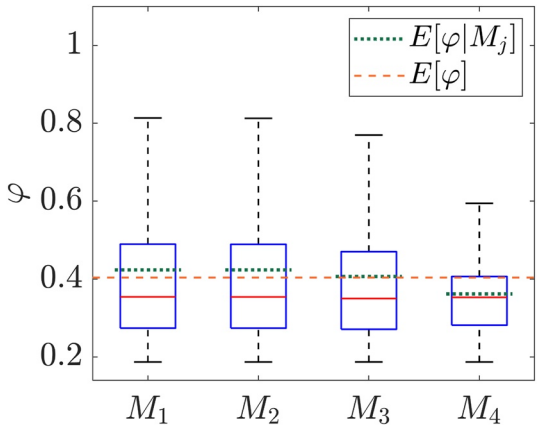

Figure 6. Boxplot representing the distributions of the Monte Carlo samples of $\varphi|M_j$, conditional to each model M_j. Results are presented together with the conditional and unconditional (Multi-Model) averages of φ (i.e., $\mathbb{E}[\varphi|M_j]$ and $\mathbb{E}[\varphi]$, respectively).

Doherty (2015a), we then consider a manual regularization scheme. Resorting to the latter stems from the impossibility of relying on expert knowledge to restrict the prior parameter space, consistent with the still limited knowledge about geochemical scenarios associated with CECs of the type we consider here. In details, we adopt the strategy described in the following.

- Parameter k_4 is set to a constant value, corresponding to the mean value of the interval listed in Table 1. This choice is consistent with the observation that available data are associated with temporal regions where process P_4 is non-influential (see Section 3.1.1).
- We choose to calibrate only r_{max3} (or r'_{max3}) among the parameters associated with process P_3. The remaining parameters are excluded from calibration and set to constant values (in models M_1, M_2, M_3) by virtue of their markedly low contributions to all *AMAE* indices (in both SM and MM contexts; see also Supplementary Material A in Supporting Information S1 and Section 3.1). Such values correspond to the mean values of the intervals listed in Table 1.

In summary, we perform model calibration upon estimating three parameters for each competing model, that is, k_1, k_2 and r_{max3} (or r'_{max3}). The results of ML calibration of all alternative models are listed in Table 2. Consistent with our expectations (based on the analysis of MM-GSA results), k_1 and k_2 are estimated with smaller uncertainty than r_{max3} (or its counterpart r'_{max3}) in all models, as seen from the values attained by the associated coefficient of variation (*CV*s; see Table 2).

3.3. Models Comparison and Identification

Here, we focus on the results associated with the evaluation of posterior model probabilities based on the Kashyap Information Criterion (KIC), as discussed in Section 2.

The results listed in Table 3 show that the minimum value of KIC in our model set is experienced for (calibrated) model M_3. Accordingly, M_3 is assigned the highest likelihood under the considered data availability. The most complex model (i.e., M_1) exhibits the lowest NLL_{MIN} among the four competing models. In this sense, the high-complexity model is associated with the highest ability to match experimental results as compared against other candidates. Nevertheless, the performance of M_1 is penalized (in terms of KIC) due to its higher complexity in processes parameterization as compared against model M_3. The latter model appears to be characterized by the best balance between complexity and accuracy. Models M_2 and M_4 are associated with low probability values. This result is consistent with the observation that model M_2 does not yield significant improvements if

Table 2
Maximum Likelihood (ML) Results of Model Calibration in Terms of Mean (i.e., Estimated Value) and Variance of Model Parameters

	$\log_{10} k_1$				$\log_{10} k_2$			
ML Results	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4
Mean	9.109	9.121	9.091	9.222	2.886	2.890	2.903	2.852
Variance	0.039	0.026	0.024	0.048	0.142	0.186	0.121	0.097
CV	0.022	0.018	0.017	0.024	0.131	0.149	0.120	0.109
	$\log_{10} r_{max3}$				$\log_{10} r'_{max3}$			
ML Results	M_1	M_2	M_3	M_4	M_1	M_2	M_3	M_4
Mean	−9.301	−9.390	-	-	-	-	−0.444	−2.208
Variance	10.530	0.854	-	-	-	-	30.990	0.314
CV	0.349	0.098	-	-	-	-	12.538	0.254

Note. Missing values correspond to specific models where processes formulation does not rely on these parameters.

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3
Minimized Negative-Log-Likelihood (NLL_{MIN}), Kashyap Information Criterion (KIC), and Posterior Probability ($\mathbb{P}$) of Each Competing Model

	M_1	M_2	M_3	M_4
NLL_{MIN}	−0.9228	−0.7809	−0.8280	−0.6611
KIC	−61.11	−58.56	−62.05	−56.81
$\mathbb{P}$	33.31%	9.33%	53.47%	3.89%

compared to model M_3, while including a larger set of model parameters. Model M_4 displays the largest value of NLL_{MIN} among the four models, and is thus penalized. These conclusions are also supported by the results of the calibration of model M_3, which can be interpreted in light of the results of MM-GSA on φ. Considering that three out of five uncertain parameters in M_3 (i.e., k_1, k_2 and k_4) are more influential on model outputs as compared against the remaining ones (r'_{max3} and K_{inhib3}), model M_3 is favored by the joint occurrence of the following features: (a) its most influential parameters are satisfactorily estimated, with the sole exception of k_4; (b) the lack of success in obtaining a satisfactory estimate of k_4 should be chiefly ascribed to the quality of the available data and not to any specific feature of M_3, consistent with the observation that this limitation cannot be circumvented even upon relying on more complex models (see Figure 4d); (c) even as the results of the calibration of model M_3 might appear penalized by the impossibility to reduce the estimation uncertainty of r'_{max3} and K_{inhib3}, the values assigned to these parameters affect only minimally model outputs (see Figure 5c); and (d) the (scarce) performance of model M_4 in the calibration stage suggests that neglecting the effects of inhibition in process P_3 markedly reduces the model ability to fit the available data (thus, the performance of models that are more simplified than M_3 tends to worsen further). Joint evaluation of all of these elements supports the conclusion that M_3 is the best candidate model in our set (under the currently available data).

4. Conclusions

In this work, we developed a strategy to simplify and calibrate over-parameterized models, which are quite common to describe complex CECs degradation paths. Our modeling strategy rests on three key methodological steps: (I) process-oriented, MM-GSA, (II) ML model calibration, and (III) model identification. Our work combines these three aspects within a unified framework and demonstrates the potential of their joint use in aquatic contamination scenarios. It then leads to the following major conclusions.

- We assess the global methodology on a highly parameterized bio-mediated transformation model of DCF. The latter has been selected as a key example of a high-complexity scenario of marked relevance in modern applications related to Emerging Contaminants. In addition to being well established, the setup we consider is of particular interest because it encompasses a considerable number of uncertain parameters while being associated with limited data availability. Thus, it reflects a situation which is commonly faced in a variety of problems across various scientific/technical fields of application. In the specific scenario illustrated in our study, we consider various model processes, these being seen as individual components in a model formulation. Values obtained for the *AMAE* indices allow quantitative ranking of the contribution of each individual process (and ensuing uncertain parameters) to the expected value of the quantities of interest. Nitrosation of DCF (i.e., process P_1) dominates the model response, while the remaining three considered processes result in smaller (yet non-negligible) impacts on the considered quantities; yet the relative impact of processes evolves with time. The interaction and relative importance of the different processes exhibit a direct link to the denitrification cycle dynamics. Our analysis allows to improve the understanding about the relative importance of all the processes involved in the DCF degradation cycle, with potential application to prototyping of new contaminant biotransformation models.
- MM-GSA indices can quantify the relative contributions ascribed to the mathematical description of the involved system processes and to the embedded parameters at various time levels. We find that for the DCF biotransformation data set here considered, available observations provide useful information content to select an appropriate model formulation and to constrain parameters of two out of four considered processes (i.e., nitrosation -P_1- and formation of Nitro-Dcf -P_2-). The parameters of the remaining two processes (formation of Aminyl-Dcf -P_3- and back-transformation -P_4-) have a negligible impact on the expected value of model outputs (as measured in terms of *AMAE*) at most time levels where data are available. This suggests that the amount of information embedded in the data set does not allow reliable estimation of these parameters. In this sense, relying on our MM-GSA can be beneficial to assist future studies for the design of experimental campaigns, with the aim of maximizing the information content carried by observations.
- The results of our MM-GSA can be ultimately used to drive parameter estimation and model selection. To this end, we estimate posterior probabilities of the competing models considered as rendered through the Kashyap Information Criterion. Accordingly, a simplified formulation of DCF biotransformation is revealed

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to be favored despite the outcome of its ML calibration seems slightly less accurate as compared against its high-complexity counterpart. The limited amount of information associated with the available data set favors a more simplified conceptualization of the system. The latter is thus considered as the best compromise between the level of model complexity and the resulting estimation uncertainty.

- The choice to apply the proposed approach to a set of candidate models obtained upon progressive simplification of a complex reference model does not constitute a limitation for the applicability of our conceptual framework. The latter is fully compatible with the use of model formulations eventually relying on mutually exclusive hypotheses on the mechanisms driving system evolution. We also remark that the selection of a batch scenario does not constitute a conceptual limitation of this study, whose theoretical framework is fully applicable to problem settings including, for example, subsurface transport processes. Computational requirements related to the implementation of a Monte Carlo analysis for reactive transport models should be taken into careful consideration. We note that the complete analysis of the scenario we consider (which is based on a set of $5 \cdot 10^4$ realizations for each candidate model to ensure convergence of all quantities of interest; not shown) requires about 140 days (overall) of computational time (approximately 1 min for each PHREEQC simulation) on a four system core-based machine with Intel(R) Core(TM) i7-7500U CPU and 16 GB RAM. When considering transport, resorting to a surrogate (or reduced complexity) model to approximate the full system model might alleviate the computational burden. Otherwise, it is noted that relying on an approximate model might introduce additional sources of uncertainties. This element needs to be carefully and rigorously considered to avoid shadowing some important results of the analysis. While surrogate models have been considered to address model parameter uncertainty (see, e.g., Dell'Oca et al., 2017), doing so in the presence of multiple models and a high number of parameters is still challenging and deserves targeted future investigations.

Data Availability Statement

Simulation data sets for this research are available at: https://doi.org/10.5281/zenodo.7513795 (Ceresa et al., 2022). Softwares for this research are publicly available at: https://www.usgs.gov/software/phreeqc-version-3 (Parkhurst & Appelo, 2021), https://pesthomepage.org/programs (Doherty, 2015b).

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

References

Appelo, C. A. J., & Postma, D. (2004). *Geochemistry, groundwater and pollution*. CRC press.

Barbieri, M., Carrera, J., Ayora, C., Sanchez-Vila, X., Licha, T., Nödler, K., 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. https://doi.org/10.1016/j.scitotenv.2012.02.058

Bouly, L., Courant, F., Bonnafé, E., Carayon, J.-L., Malgouyres, J.-M., Vignet, C., et al. (2022). Long-term exposure to environmental diclofenac concentrations impairs growth and induces molecular changes in lymnaea stagnalis freshwater snails. *Chemosphere*, *291*, 133065. https://doi.org/10.1016/j.chemosphere.2021.133065

Canelles, A., Rodríguez-Escales, P., Modrzyński, J. J., Albers, C., & Sanchez-Vila, X. (2021). Impact of compost reactive layer on hydraulic transport and C & N cycles: Biogeochemical modeling of infiltration column experiments. *Science of the Total Environment*, *770*, 145490. https://doi.org/10.1016/j.scitotenv.2021.145490

Carrera, J., & Neuman, S. P. (1986). Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resources Research*, *22*(2), 199–210. https://doi.org/10.1029/WR022i002p00199

Ceresa, L., Guadagnini, A., Porta, G. M., & Riva, M. (2021). Formulation and probabilistic assessment of reversible biodegradation pathway of diclofenac in groundwater. *Water Research*, *204*, 117466. https://doi.org/10.1016/j.watres.2021.117466

Ceresa, L., Guadagnini, A., Rodríguez-Escales, P., Riva, M., Porta, G. M., & Sanchez-Vila, X. (2022). Datasets for numerical Monte Carlo simulations of diclofenac bio-degradation in a soil-water system [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.7513795

Ceriotti, G., Guadagnini, L., Porta, G., & Guadagnini, A. (2018). Local and global sensitivity analysis of Cr (VI) geogenic leakage under uncertain environmental conditions. *Water Resources Research*, *54*(8), 5785–5802. https://doi.org/10.1029/2018WR022857

Chen, Y.-C., & Ma, H.-W. (2006). Model comparison for risk assessment: A case study of contaminated groundwater. *Chemosphere*, *63*(5), 751–761. https://doi.org/10.1016/j.chemosphere.2005.08.011

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. https://doi.org/10.1016/j.scitotenv.2016.05.048

Dai, H., Ye, M., Walker, A. P., & Chen, X. (2017). A new process sensitivity index to identify important system processes under process model and parametric uncertainty. *Water Resources Research*, *53*(4), 3476–3490. https://doi.org/10.1002/2016WR019715

Dell'Oca, A., Riva, M., & Guadagnini, A. (2017). Moment-based metrics for global sensitivity analysis of hydrological systems. *Hydrology and Earth System Sciences*, *21*(12), 6219–6234. https://doi.org/10.5194/hess-21-6219-2017

Dell'Oca, A., Riva, M., & Guadagnini, A. (2020). Global sensitivity analysis for multiple interpretive models with uncertain parameters. *Water Resources Research*, *56*(2), e2019WR025754. https://doi.org/10.1029/2019WR025754

Doherty, J. (2015a). *Calibration and uncertainty analysis for complex environmental models*. Watermark Numerical Computing Brisbane.

Doherty, J. (2015b). *PEST, version 17* [Software]. Retrieved from https://pesthomepage.org/programs

Elgendy, A. M., & Porta, G. M. (2021). Impact of reservoir geochemistry on low salinity waterflooding: Global sensitivity analysis. *Journal of Petroleum Science and Engineering*, *207*, 109056. https://doi.org/10.1016/j.petrol.2021.109056

Farhat, S. K., Newell, C. J., Lee, S. A., Looney, B. B., & Falta, R. W. (2022). Impact of matrix diffusion on the migration of groundwater plumes for perfluoroalkyl acids (PFAAs) and other non-degradable compounds. *Journal of Contaminant Hydrology*, *247*, 103987. https://doi.org/10.1016/j.jconhyd.2022.103987

Fent, K., Weston, A. A., & Caminada, D. (2006). Ecotoxicology of human pharmaceuticals. *Aquatic Toxicology*, *76*(2), 122–159. https://doi.org/10.1016/j.aquatox.2005.09.009

Greskowiak, J., Prommer, H., Massmann, G., & Nützmann, G. (2006). Modeling seasonal redox dynamics and the corresponding fate of the pharmaceutical residue phenazone during artificial recharge of groundwater. *Environmental Science & Technology*, *40*(21), 6615–6621. https://doi.org/10.1021/es052506t

Hauck, M., Huijbregts, M. A., Armitage, J. M., Cousins, I. T., Ragas, A. M., & van de Meent, D. (2008). Model and input uncertainty in multimedia fate modeling: Benzo [a] pyrene concentrations in Europe. *Chemosphere*, *72*(6), 959–967. https://doi.org/10.1016/j.chemosphere.2008.03.014

Heberer, T., & Adam, M. (2004). Transport and attenuation of pharmaceutical residues during artificial groundwater replenishment. *Environmental Chemistry*, *1*(1), 22–25. https://doi.org/10.1071/EN04008

Im, J., Rizzo, C. B., & de Barros, F. P. (2020). Resilience of groundwater systems in the presence of bisphenol a under uncertainty. *Science of the Total Environment*, *727*, 138363. https://doi.org/10.1016/j.scitotenv.2020.138363

Janetti, E. B., Dror, I., Guadagnini, A., Riva, M., & Berkowitz, B. (2012). Estimation of single-metal and competitive sorption isotherms through maximum likelihood and model quality criteria. *Soil Science Society of America Journal*, *76*(4), 1229–1245. https://doi.org/10.2136/sssaj2012.0010

Jurado, A., Walther, M., & Díaz-Cruz, M. S. (2019). Occurrence, fate and environmental risk assessment of the organic microcontaminants included in the watch lists set by EU decisions 2015/495 and 2018/840 in the groundwater of Spain. *Science of the Total Environment*, *663*, 285–296. https://doi.org/10.1016/j.scitotenv.2019.01.270

Kashyap, R. L. (1982). Optimal choice of AR and MA parts in autoregressive moving average models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2), 99–104. https://doi.org/10.1109/TPAMI.1982.4767213

Kiecak, A., Sassine, L., Boy-Roura, M., Elsner, M., Mas-Pla, J., La Salle, C. L. G., & Stumpp, C. (2019). Sorption properties and behaviour at laboratory scale of selected pharmaceuticals using batch experiments. *Journal of Contaminant Hydrology*, *225*, 103500. https://doi.org/10.1016/j.jconhyd.2019.103500

Kumar, A., Chang, B., & Xagoraraki, I. (2010). Human health risk assessment of pharmaceuticals in water: Issues and challenges ahead. *International Journal of Environmental Research and Public Health*, *7*(11), 3929–3953. https://doi.org/10.3390/ijerph7113929

La Farre, M., Pérez, S., Kantiani, L., & Barceló, D. (2008). Fate and toxicity of emerging pollutants, their metabolites and transformation products in the aquatic environment. *TrAC Trends in Analytical Chemistry*, *27*(11), 991–1007. https://doi.org/10.1016/j.trac.2008.09.010

Lonappan, L., Brar, S. K., Das, R. K., Verma, M., & Surampalli, R. Y. (2016). Diclofenac and its transformation products: Environmental occurrence and toxicity-a review. *Environment International*, *96*, 127–138. https://doi.org/10.1016/j.envint.2016.09.014

Nödler, K., Licha, T., Barbieri, M., & Pérez, S. (2012). Evidence for the microbially mediated abiotic formation of reversible and non-reversible sulfamethoxazole transformation products during denitrification. *Water Research*, *46*(7), 2131–2139. https://doi.org/10.1016/j.watres.2012.01.028

Osorio, V., Sanchís, J., Abad, J. L., Ginebreda, A., Farré, M., Pérez, S., & Barceló, D. (2016). Investigating the formation and toxicity of nitrogen transformation products of diclofenac and sulfamethoxazole in wastewater treatment plants. *Journal of Hazardous Materials*, *309*, 157–164. https://doi.org/10.1016/j.jhazmat.2016.02.013

Parkhurst, D., & Appelo, T. (2021). *PHREEQC, version 3.6.*2 [Software]. Retrieved from https://www.usgs.gov/software/phreeqc-version-3

Parkhurst, D. L., & Appelo, C. A. J. (2013). Description of input and examples for phreeqc version 3—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. *US geological survey techniques and methods*, *6*(A43), 497.

Reberski, J. L., Terzić, J., Maurice, L. D., & Lapworth, D. J. (2022). Emerging organic contaminants in karst groundwater: A global level assessment. *Journal of Hydrology*, *604*, 127242. https://doi.org/10.1016/j.jhydrol.2021.127242

Rodríguez-Escales, P., Fernàndez-Garcia, D., Drechsel, J., Folch, A., & Sanchez-Vila, X. (2017). Improving degradation of emerging organic compounds by applying chaotic advection in m anaged a quifer r echarge in randomly heterogeneous porous media. *Water Resources Research*, *53*(5), 4376–4392. https://doi.org/10.1002/2016wr020333

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. https://doi.org/10.1016/j.watres.2016.09.034

Rozman, D., Hrkal, Z., Eckhardt, P., Novotná, E., & Boukalová, Z. (2015). Pharmaceuticals in groundwaters: A case study of the psychiatric hospital at Horní Beřkovice, Czech republic. *Environmental Earth Sciences*, *73*(7), 3775–3784. https://doi.org/10.1007/s12665-014-3663-1

Ruggeri, B. (2009). Chemicals exposure: Scoring procedure and uncertainty propagation in scenario selection for risk analysis. *Chemosphere*, *77*(3), 330–338. https://doi.org/10.1016/j.chemosphere.2009.07.014

Scheytt, T. J., Mersmann, P., & Heberer, T. (2006). Mobility of pharmaceuticals carbamazepine, diclofenac, ibuprofen, and propyphenazone in miscible-displacement experiments. *Journal of Contaminant Hydrology*, *83*(1–2), 53–69. https://doi.org/10.1016/j.jconhyd.2005.11.002

Schimmelpfennig, S., Kirillin, G., Engelhardt, C., Dünnbier, U., & Nützmann, G. (2016). Fate of pharmaceutical micro-pollutants in Lake Tegel (Berlin, Germany): The impact of lake-specific mechanisms. *Environmental Earth Sciences*, *75*(10), 1–12. https://doi.org/10.1007/s12665-016-5676-4

Sobol, I. M. (1993). Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling in Civil Engineering*, *1*, 407–414.

Sobol, I. M. (1998). On quasi-Monte Carlo integrations. *Mathematics and Computers in Simulation*, *47*(2–5), 103–112. https://doi.org/10.1016/S0378-4754(98)00096-2

Walker, A. P., Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Asao, S., Hickler, T., et al. (2015). Predicting long-term carbon sequestration in response to co2 enrichment: How and why do current ecosystem models differ? *Global Biogeochemical Cycles*, *29*(4), 476–495. https://doi.org/10.1002/2014GB004995

Ye, M., Lu, D., Neuman, S. P., & Meyer, P. D. (2010). Comment on "Inverse Groundwater Modeling for Hydraulic Conductivity Estimation Using Bayesian Model Averaging and Variance Window" by Frank T.-C. Tsai and Xiaobao Li. *Water Resources Research*, *46*(2), 2802. https://doi.org/10.1029/2009WR008501

Ye, M., Meyer, P. D., & Neuman, S. P. (2008). On model selection criteria in multimodel analysis. *Water Resources Research*, *44*(3). https://doi.org/10.1029/2008WR006803

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Zhang, X., Niu, G.-Y., Elshall, A. S., Ye, M., Barron-Gafford, G. A., & Pavao-Zuckerman, M. (2014). Assessing five evolving microbial enzyme models against field measurements from a semiarid savannah—What are the mechanisms of soil respiration pulses? *Geophysical Research Letters*, *41*(18), 6428–6434. https://doi.org/10.1002/2014GL061399

Zhang, Y., Sun, K., Li, Z., Chai, X., Fu, X., Kholodkevich, S., et al. (2021). Effects of acute diclofenac exposure on intestinal histology, antioxidant defense, and microbiota in freshwater crayfish (Procambarus clarkii). *Chemosphere*, *263*, 128130. https://doi.org/10.1016/j.chemosphere.2020.128130

19447973, 2023, 1, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022WR033183 by CochraneItalia, Wiley Online Library on [20/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License