Bayesian inversion of bacterial physiology and dissolved organic carbon biodegradability on water incubation data

Shuaitao Wang, Nicolas Flipo, Josette Garnier, Thomas Romary

Science of the Total Environment 955 (2024) 177252

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Bayesian inversion of bacterial physiology and dissolved organic carbon biodegradability on water incubation data

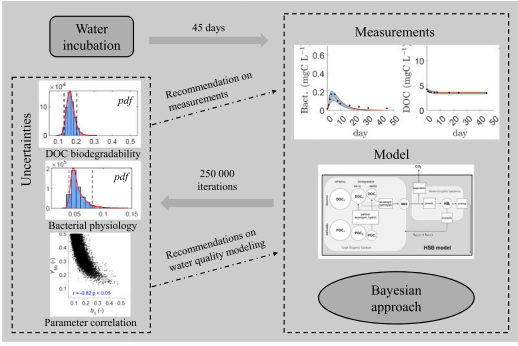
Shuaitao Wang [a,c,*], Nicolas Flipo [b], Josette Garnier [a], Thomas Romary [b]

[a] Sorbonne Université, CNRS, EPHE, UMR METIS, Paris 75005, France
[b] Mines Paris, PSL University, Center for geosciences and geoengineering, Fontainebleau 77300, France
[c] BRGM, Geological Survey of France, Orléans 45100, France

HIGHLIGHTS

- A multiple MCMC method is used for the first time to interpret water incubation data.
- Inferred bacterial physiology is consist with older laboratory experiments.
- Inferred DOC biodegradability allows for the discussion of a new model structure.
- Uncertainties in model parameters and measurement errors are investigated.
- Measuring bacterial information during water incubation is recommended.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Jose Julio Ortega-Calvo

Keywords:
DOC biodegradability
Bacterial physiology
Aquatic biogeochemical modeling
Bayesian inversion
Parameter uncertainties

ABSTRACT

In aquatic ecosystems, dissolved organic carbon (DOC) plays a significant role in the global carbon cycle. Microorganisms mineralize biodegradable DOC, releasing greenhouse gases (carbon dioxide, methane) into the atmosphere. Extensive research has focused on the concentrations and biodegradability of DOC in aquatic systems worldwide. However, little attention has been given to uncertainties regarding the physiological characteristics of heterotrophic bacteria, which are crucial for biogeochemical modeling. In this study, the physiological properties of heterotrophic bacteria and the properties of DOC biodegradability in water are inferred through a Bayesian inversion approach. To achieve this, treated and natural water samples collected from the Seine River basin, were inoculated and incubated in laboratory. During incubation, the concentrations of DOC and heterotrophic bacteria biomass were measured. Then, a multiple Monte Carlo Markov Chains method and the HSB model (High-weight polymers, Substrate, heterotrophic Bacteria) are applied on the water incubation data. The results indicate a higher biodegradable fraction of DOC in natural water compared to treated water and significant variability in the fraction of fast biodegradable DOC within 5 days in both water samples. The significant variability highlights the uncertainties/challenges in the HSB model parameterization. The seven water samples used in the paper serve as a proof of concept. They are from various origins and display the potential of the method to identify parameter values in a large range of values. Because mortality rate of heterotrophic bacteria at 20 °C (k_{d20}) showed a remarkable stability at 0.013 h^{-1}, we considered that this parameter can be fixed at this

* Corresponding author at: Sorbonne Université, CNRS, EPHE, UMR METIS, Paris 75005, France.
E-mail addresses: shuaitao.wang@sorbonne-universite.fr, s.wang@brgm.fr (S. Wang).

https://doi.org/10.1016/j.scitotenv.2024.177252
Received 30 July 2024; Received in revised form 5 October 2024; Accepted 25 October 2024
Available online 3 November 2024

value. The maximum growth rates at 20 °C (μ_{max20}) was 0.061 h^{-1} while optimal growth yield (Y) estimated at 0.34 for treated water and at 0.25 for natural water. All these parameter values are well in accordance with previous determinations.

1. Introduction

The dissolved organic carbon (DOC) present in surface aquatic ecosystems serves a pivotal role in the global carbon (C) cycle (Cole et al., 2007; Battin et al., 2009; Butman and Raymond, 2011; Raymond et al., 2013; Drake et al., 2018), as well as in the estimation of aquatic ecosystem metabolism (Vilmin et al., 2016; Battin et al., 2023). The remineralization process of organic carbon releases greenhouse gases such as carbon dioxide and methane into the atmosphere (Hotchkiss et al., 2015; Deemer et al., 2016; Stanley et al., 2016; Prairie et al., 2018; Marescaux et al., 2020; Hao et al., 2021; Yan et al., 2022). A recent study has indicated that a significant portion, approximately 37 %, of terrestrial organic carbon is respired by rivers (Battin et al., 2023). Biodegradable DOC (BDOC), which can be decomposed by microorganisms, emerges as a crucial variable for understanding the global C cycle (Vonk et al., 2015; Liu and Wang, 2022).

Servais et al. (1987) developed a method to determine the biodegradable fraction of DOC in waters through dark incubation, which relies on the microbial degradation-induced loss of DOC. This method stands as the most widely employed and accurate technique to date (Begum et al., 2023). Research efforts have extensively explored the concentration, composition, biodegradability, and sources of DOC in aquatic systems (Søndergaard and Worm, 2001; Wickland et al., 2007; Holmes et al., 2008; Fellman et al., 2008; Wickland et al., 2012; Abbott et al., 2014; Vonk et al., 2015; Goffin et al., 2017; Begum et al., 2019), with recent comprehensive reviews available (Fellman et al., 2010; Li and Hur, 2017; Liu and Wang, 2022; Begum et al., 2023). These studies primarily focus on the variations in DOC concentration and biodegradability across aquatic ecosystems worldwide, along with investigations into the key drivers influencing DOC concentrations and biodegradability (Larouche et al., 2015; O'Donnell et al., 2016; Mutschlecner et al., 2018; Liu et al., 2021).

However, the concurrent physiological properties of heterotrophic bacteria during incubation experiments, including growth rate, mortality rate, and growth yield, have been scarcely explored. These properties are indispensable for modeling the biogeochemical functioning of aquatic ecosystems, and their uncertainties are of significant interest to water quality modelers, biogeochemists, and environmental engineers. The HSB (High-weight polymers, Substrate, heterotrophic Bacteria) mechanistic model (Billen and Servais, 1989; Billen, 1991; Wang et al., 2024) facilitates the simulation of organic carbon degradation by microorganisms, wherein the growth of heterotrophic bacteria and the hydrolysis of organic carbon (both dissolved and particulate) are represented by Monod-Michaelis-Menten kinetic equations (Monod, 1949; Michaelis and Menten, 1913). Since the initiation of the PIREN-Seine program (https://www.piren-seine.fr/) in 1989, the parameter values of heterotrophic bacteria physiological processes in the HSB model, including maximum growth rate, mortality rate, growth yield, and biodegradability of organic carbon, have been determined at 20 °C through laboratory experiments conducted on the Seine River (Servais et al., 1985; Garnier et al., 1992a, 1992b; Bariller and Garnier, 1993). These parameter values, coupled with the modeling platforms of the PIREN-Seine program (pyNuts-Riverstrahler, ProSe, and Barman), have facilitated numerous applications aiming at understanding the biogeochemical functioning of river systems (Billen et al., 1994; Garnier et al., 1995; Even et al., 1998, 2004; Flipo et al., 2004, 2007; Even et al., 2007; Thieu et al., 2009, 2010; Le et al., 2015; Vilmin et al., 2016; Garnier et al., 2018; Romero et al., 2019; Marescaux et al., 2020) and reservoirs (Garnier and Billen, 1994; Garnier et al., 2000; Yan et al., 2022).

However, in these applications, the parameter values remain constant during the simulation period, and their uncertainties remain unclear. Studies focusing on sensitivity analyses of the C-RIVE model, which integrates the HSB model, have identified the most influential parameters on dissolved oxygen (DO) under various conditions (Wang et al., 2018; Hasanyar et al., 2023b). These results have facilitated the coupling of the deterministic water quality model ProSe (Even et al., 1998) with data assimilation methods to create ProSe-PA program (Wang et al., 2019, 2023, https://gitlab.com/prose-pa/prose-pa). Data assimilation enables the inference of physiological properties of living communities using measured DO concentrations, thereby enhancing the simulation of Seine River metabolism (Wang et al., 2022). Recent efforts have attempted to quantify the DOC biodegradability of different water inflows (rivers, combined sewer overflows, wastewater treatment plant discharges) by assimilating observed DO data (Hasanyar et al., 2023a). One of the recommendations proposed by Hasanyar et al. (2023a) is to establish a data assimilation system capable of identifying a plurality of DOC biodegradability based on its origin (rivers, combined sewer overflows, or discharges from wastewater treatment plants). The composition and sources of DOC influence the variations in DOC biodegradability (Liu et al., 2021; Liu and Wang, 2022; Begum et al., 2023). Thus, it is imperative to accurately characterize the properties of DOC biodegradability and joint properties of heterotrophic bacteria in different water inflows.

In this study, seven water samples were collected from the Seine River basin. These water samples were sourced from surface water (Seine River, standing water), groundwater, or treated water from wastewater treatment plants (WWTPs) with varying nominal capacities. The objective of this study is to infer the physiological properties of heterotrophic bacteria and the properties of DOC biodegradability in the seven water samples through a Bayesian inversion approach. To achieve this, the water samples were inoculated and incubated in the laboratory. During incubation, the concentrations of DOC and heterotrophic bacteria biomass were measured using wet oxidation and epifluorescence microscopy, respectively (Garnier et al., 2021). Then, a multiple Markov Chain Monte Carlo (MCMC) method and the HSB model are applied on the water incubation data. The uncertainties associated with these properties, their correlations, and measurement errors are subsequently discussed. The results provide valuable insights and pragmatic recommendations on aquatic biogeochemical modeling. The study demonstrates the effectiveness of the Bayesian approach and the HSB model. This methodology can be extended to analyze additional water incubation datasets in future research or applied to previous samples with existing data, enabling a deeper understanding of the influencing factors of DOC/BDOC and the uncertainties in model parameters.

2. Material and methods

2.1. Water incubation experiment

2.1.1. Water samples

Seven water samples were collected from the Seine River basin (Table 1). Among these, three samples were drawn from the effluent of three WWTPs with different nominal capacities during February 2021. They are designated as Boissettes, Rosny, and St-Thibault respectively (Table 1). Two additional samples were procured from standing water, known as "Grande Mare", i.e. a pond, situated in the Orgeval basin, in February and May 2021. They are named as Pond-Feb. and Pond-May respectively (Table 1). The remaining two samples were sourced from the surface water and alluvial deposits of the Seine River within the

Table 1
Water samples collected from different sites in Seine River basin.

Sample name	Sample type	Sample time	Capacity (PE)*	DOC (mgC L^{-1})
Boissettes	WWTP (treated water)	Feb - 2021	80,000	7.00
Rosny	WWTP (treated water)	Feb - 2021	135,500	6.87
St-Thibault	WWTP (treated water)	Feb - 2021	400,000	8.10
Alluvial	Hydrosystem (groundwater)	Oct - 2021	–	4.31
Pond-Feb.	Hydrosystem (surface water)	Feb - 2021	–	4.26
Pond-May	Hydrosystem (surface water)	May - 2021	–	12.00
River	Hydrosystem (surface water)	Oct - 2021	–	4.70

* PE: population equivalence

designated Bassée region of France during October 2021. They are identified as River and Alluvial respectively (Table 1).

2.1.2. Water incubation in laboratory

The seven water samples were first sieved through a calcinated Whatman GF/F filter (45 mm diameter, 0.75 μm porosity). Then, the filtered water samples (2 l) were inoculated with 0.5 % volume of the indigenous water type and incubated independently at an ambient temperature of 20 °C in the dark and under agitation for 45 days. During the length of each batch experiment (7 in total), one water sample was taken 12 times during the 45 days of incubation. The concentrations of DOC and bacterial biomass were measured by Aurora 1030 W TOC Analyzer (Garnier et al., 2021) and epifluorescence microscopy (Garnier et al., 1992a) respectively, all in duplicate and averaged. This incubation method was developed by Servais et al. (1995) for determining the biodegradable fraction of DOC in waters. The biodegradable fraction of DOC is assumed to be fully degraded through bacterial activities after 45 days.

2.2. Mathematical modeling of water incubation experiment

2.2.1. HSB model

To simulate the water incubation experiments, the HSB model developed by Billen et al. (1994) is employed. The HSB model explicitly captures the mechanisms underlying organic carbon degradation and heterotrophic bacterial activity, including growth, mortality, and respiration. The model delineates three key variables (depicted in Fig. 1): **H**, representing high-weight polymers encompassing dissolved and particulate organic carbon (DOC and POC); **S**, denoting small monomeric substrate (SMS), which emerges as a hydrolysis product of high-weight polymers and serves as a direct energy source for the growth and respiration of heterotrophic bacteria (H**B**).

2.2.2. Organic carbon in HSB model

The total organic carbon is conceptually partitioned into dissolved (DOC) and particulate (POC) phases, each further subdivided into three pools associated with specific biodegradability: (1) rapidly biodegradable within a 5-day period (DOC_1 and POC_1); (2) slowly biodegradable within a 45-day period (DOC_2 and POC_2); and (3) refractory (DOC_3 and POC_3). Consequently, the HSB model encompasses seven distinct pools of organic carbon (DOC_1, DOC_2, DOC_3, POC_1, POC_2, POC_3, and SMS).

This paper places particular emphasis on investigating the properties of DOC biodegradability and heterotrophic bacteria. Hasanyar et al. (2023b), utilizing observed dissolved oxygen (DO) concentrations, introduced two parameters, denoted as b_1 and s_1, to capture the relationship between DOC and its three constituent pools (Eqs. (1), (2), and (3)). Here, b_1 signifies the fraction of biodegradable DOC, while $s_1 \times b_1$ quantifies the fraction of fast biodegradable DOC (DOC_1).

$$DOC_1 = s_1 \times b_1 \times DOC \quad (1)$$

$$DOC_2 = (1 - s_1) \times b_1 \times DOC \quad (2)$$

$$DOC_3 = (1 - b_1) \times DOC \quad (3)$$

2.2.3. Dynamics of heterotrophic bacteria in HSB model

The dynamics of heterotrophic bacteria encompass growth, mortal-

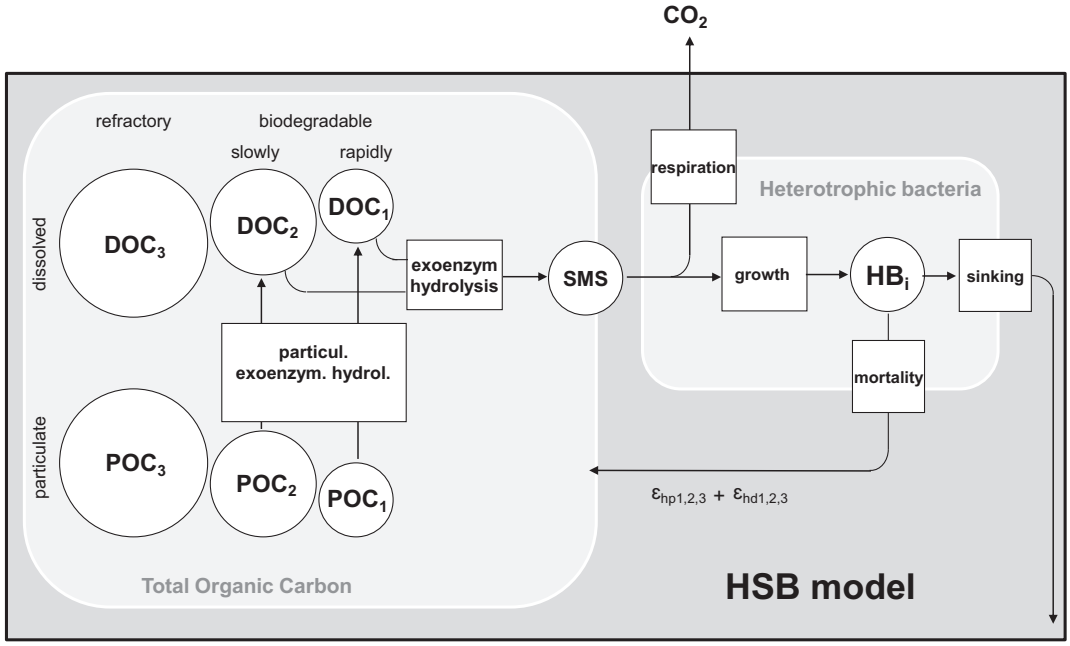


Fig. 1. Flowchart of HSB model. DOC_1: rapidly biodegradable dissolved organic carbon; DOC_2: slowly biodegradable dissolved organic carbon; DOC_3: refractory dissolved organic carbon; POC_1: rapidly biodegradable particulate organic carbon; POC_2: slowly biodegradable particulate organic carbon; POC_3: refractory particulate organic carbon; SMS: small monomeric substrate; HB: heterotrophic bacteria.

ity, and respiration processes. Within the HSB model, growth is represented by a Monod function (Monod, 1949), where a maximum growth rate at 20 °C ($\mu_{max20,hb}$, [h^{-1}], Eq. (4)) is employed. Mortality is described by a first-order kinetics formulation with a mortality rate at 20 °C (k_{d20}, [h^{-1}]). In practical applications, parameter values at temperature T °C are derived by adjusting the values determined at 20 °C using a temperature function (Billen et al., 1994; Wang et al., 2024). Given that water samples were incubated at 20 °C, Eqs. (4) and (5) do not incorporate any temperature effects.

$$\mu_{hb} = \mu_{max20,hb}\frac{[\mathrm{SMS}]}{[\mathrm{SMS}] + K_{sms,hb}} \tag{4}$$

$$\frac{d[\mathrm{HB}]}{dt} = \left(\mu_{hb} - k_{d20,hb}\right)[\mathrm{HB}] \tag{5}$$

With $\mu_{max20,hb}$: maximum growth rate of heterotrophic bacteria, [h^{-1}].

[SMS]: concentration of small monomeric substrate, [mgC L^{-1}].

$K_{sms,hb}$: half-saturation constant for SMS of heterotrophic bacteria, [mgC L^{-1}].

$k_{d20,hb}$: mortality rate of heterotrophic bacteria, [h^{-1}].

[HB]: concentration of heterotrophic bacteria, [mgC L^{-1}].

The growth of bacteria depends on the availability of small monomeric substrate. The uptake of small monomeric substrate (*uptS*) by heterotrophic bacteria is determined through the utilization of a growth yield (Y_{hb}) which has been found constant in the 8–25°C range of temperature (Bariller and Garnier, 1993).

$$uptS = \frac{\mu_{hb}}{Y_{hb}}[\mathrm{HB}] \tag{6}$$

With μ_{hb}: effective growth rate of heterotrophic bacteria, Eq. (4), [h^{-1}].

Y_{hb}: growth yield of heterotrophic bacteria, [−].

[HB]: concentration of heterotrophic bacteria, [mgC L^{-1}].

uptS: uptake of small monomeric substrate for bacterial activity, [mgC L^{-1} h^{-1}].

2.2.4. Hydrolysis of organic carbon in HSB model

As previously mentioned, small monomeric substrate (SMS) is directly assimilated by heterotrophic bacteria. Consequently, high-weight polymers, represented by dissolved organic carbon (DOC) and particulate organic carbon (POC), require hydrolysis to generate SMS. Biodegradable POC (POC_1 and POC_2) undergo hydrolysis, producing biodegradable DOC (DOC_1 and DOC_2). The kinetics governing the hydrolysis of biodegradable POC in the HSB model are characterized by a first-order equation incorporating a hydrolysis rate constant (k_{poc_i}, [h^{-1}]). Conversely, DOC hydrolysis is exoenzymatic, resulting in the formation of SMS. This exoenzymatic process is mathematically described using a Michaelis-Menten function, as outlined by Michaelis and Menten (1913).

$$hyd_{doc_i} = e_{max,doc_i,hb}\frac{[\mathrm{DOC}_i]}{[\mathrm{DOC}_i] + K_{doc_i,hb}}[\mathrm{HB}] \tag{7}$$

With $e_{max,doc_i,hb}$: maximum hydrolysis rate of DOC_i for heterotrophic bacteria, $i = 1, 2$, [h^{-1}].

$K_{doc_i,hb}$: half-saturation constant for DOC_i of heterotrophic bacteria, $i = 1, 2$, [mgC L^{-1}].

[HB]: concentration of heterotrophic bacteria, [mgC L^{-1}].

2.2.5. Bacterial physiological properties and DOC biodegradability in HSB model

The physiological properties of heterotrophic bacteria within the HSB model are represented by parameters such as the maximum growth rate ($\mu_{max20,hb}$), the growth yield (Y_{hb}), and the mortality rate ($k_{d20,hb}$), all determined at 20 °C. Studies have illustrated that these parameters have a considerable influence on river metabolism (Wang et al., 2018; Hasanyar et al., 2023b). Similarly, the properties of DOC biodegradability are delineated by the parameters b_1 and s_1 (Eqs. (1) and (2)). These parameter values are crucial for aquatic biogeochemical modeling and their uncertainties are of great interest to scientific community in biogeochemical modeling.

In this study, the three parameters representing heterotrophic bacteria properties ($\mu_{max20,hb}$, Y_{hb}, and $k_{d20,hb}$), the two parameters characterizing DOC biodegradability (b_1 and s_1), and the initial concentrations of heterotrophic bacteria and DOC (hb_{init} and DOC_{init}) are selected and included in the framework of Bayesian inversion.

2.3. Framework of Bayesian inversion

2.3.1. Forward model

The HSB model simulates the dynamics of DOC and heterotrophic bacteria, given the above mentioned parameters and the initial concentrations. However, these parameters are considered unknown and the initial concentrations are uncertain. These inputs need to be inferred from the noisy water incubation data. The Bayesian inversion method is employed to estimate these parameters under uncertainty by integrating the HSB model with the observation data. Formally, given $\mathbf{x}$, the parameter vector we seek to identify, and an initial state $\mathbf{y}_0$ (concentrations of DOC and heterotorophic bacteria), running the HSB model, denoted by M, yields a sequence $\mathbf{y}_{1:T}$ of DOC concentration and bacteria biomass along time. The problem is then to retrieve the parameters as well as the initial conditions, given the observation data $\mathbf{y}_\mathbf{t}^*$, where the subscript $\mathbf{t}$ corresponds to the times when the samples were analyzed during the incubation. We can formalize the forward model and the observation process as follows:

$$\mathbf{y}_{1:T} = M(\mathbf{y}_0, \mathbf{x}) \tag{8}$$

$$\mathbf{y}_\mathbf{t}^* = \mathbf{H}\mathbf{y}_{1:T} + \eta_\mathbf{t} \tag{9}$$

where $\mathbf{H}$ is the observation operator and $\eta_\mathbf{t}$ is the measurement error. $\mathbf{t}$ is the vector of times at which the observations were measured. Considering the stochastic unknowns $\mathbf{y_0}$, $\mathbf{x}$ and errors, we can derive their conditional probability distribution with respect to the observations. This quantity is termed the posterior distribution and is described below.

2.3.2. Bayesian inference

From Eqs. (8) and (9), we read that $\mathbf{y}_\mathbf{t}^*$ is conditionally independent of $\mathbf{y_0}$, $\mathbf{x}$, knowing $\mathbf{y}_{1:T}$. Applying the Bayes theorem (Bayes, 1763), the posterior (target) distribution of the model parameters and initial state, $f(\mathbf{x}, \mathbf{y}_0|\mathbf{y}_t^*)$, can be derived up to a normalizing constant:

$$f(\mathbf{x}, \mathbf{y}_0|\mathbf{y}_\mathbf{t}^*) \propto f(\mathbf{y}_\mathbf{t}^*|\mathbf{y}_{1:T}) f(\mathbf{y}_{1:T}|\mathbf{x}, \mathbf{y}_0) f(\mathbf{x}) f(\mathbf{y}_0) \tag{10}$$

where $f(\mathbf{x})$ and $f(\mathbf{y}_0)$ are the prior distribution of the model parameters $\mathbf{x}$ and initial states $\mathbf{y}_0$ (here, initial concentrations of DOC and heterotrophic bacteria) that are considered independent. $f(\mathbf{y}_\mathbf{t}^*|\mathbf{y}_{1:T})$ is the likelihood which gives the probability to observe $\mathbf{y}_\mathbf{t}^*$ given model states $\mathbf{y}_{1:T}$.

2.3.3. MCMC methods: From random walk Metropolis to DREAM

Although no analytical expression is available for the posterior distribution ($f(\mathbf{x}, \mathbf{y}_0|\mathbf{y}_\mathbf{t}^*)$), we can access it through sampling. This can be effectively achieved through Markov Chain Monte Carlo (MCMC) simulation (Robert and Casella, 2004). This methodology involves constructing a Markov chain that will converge in distribution toward the target distribution $f(\mathbf{x}, \mathbf{y}_0|\mathbf{y}_\mathbf{t}^*)$. The pioneering MCMC method, introduced by Metropolis et al. (1953), is commonly known as the Random Walk Metropolis (RWM) algorithm. To explore the posterior distribution, the algorithm employs trial moves (random walks or other)

from the current value of the parameters ($\mathbf{x}$) to propose new values ($\mathbf{x}^p$), where $\mathbf{x}^p \sim q(\cdot|\mathbf{x})$ and $q(\cdot)$ represents the proposal distribution. In order to ensure the convergence of the Markov chain (Hastings, 1970), probabilities for accepting the trial move from $\mathbf{x}$ to $\mathbf{x}^p$ ($\alpha_{\mathbf{x}\to\mathbf{x}^p}$) and from $\mathbf{x}^p$ to $\mathbf{x}$ ($\alpha_{\mathbf{x}^p\to\mathbf{x}}$) are introduced:

$$\alpha_{\mathbf{x}\to\mathbf{x}^p} = min\left[1, \frac{f(\mathbf{x}^p, \mathbf{y}_0^p|\mathbf{y}_t^*)q(\mathbf{x}|\mathbf{x}^p)}{f(\mathbf{x}, \mathbf{y}_0|\mathbf{y}_t^*)q(\mathbf{x}^p|\mathbf{x})}\right] \tag{11}$$

By defining a proposal distribution $q(\cdot)$ and utilizing the acceptance probability $\alpha_{\mathbf{x}\to\mathbf{x}^p}$ (Eq. (11)), we can iteratively determine whether to accept or reject trial moves, thereby approximating the target distribution. This approach is known as the Metropolis-Hastings algorithm (Hastings, 1970). In cases where the proposal distribution is symmetrical ($q(\mathbf{x}|\mathbf{x}^p) = q(\mathbf{x}^p|\mathbf{x})$), such as with a centered normal distribution, the acceptance probability $\alpha_{\mathbf{x}\to\mathbf{x}^p}$ can be simplified (Eq. (12)):

$$\alpha_{\mathbf{x}\to\mathbf{x}^p} = min\left[1, \frac{f(\mathbf{x}^p, \mathbf{y}_0^p|\mathbf{y}_t^*)}{f(\mathbf{x}, \mathbf{y}_0|\mathbf{y}_t^*)}\right] \tag{12}$$

Since the inception of the RWM algorithm (Metropolis et al., 1953), more sophisticated MCMC methods have emerged to enhance the efficiency of MCMC simulations. These advancements include both single and multiple chain methods. Single chain methods, such as adaptive Metropolis (AM) (Haario et al., 1999) and delayed rejection adaptive Metropolis (DRAM) (Haario et al., 2006), dynamically adjust the covariance of the proposal distribution. Multiple chain methods utilize differential evolution (Storn and Price, 1997; Price et al., 2005) and subspace sampling (Vrugt et al., 2009) to expedite convergence to the target distribution, exemplified by the Differential Evolution Markov chain (DE-MC) (Braak, 2006) and the DiffeRential Evolution Adaptive Metropolis (DREAM) (Vrugt et al., 2009; Vrugt, 2016). When the target distribution is hard to reach, for instance when its support is small in the space of the parameters, other interacting schemes targeting tempered versions of the posterior distribution can also be used (Romary, 2010; Bottero et al., 2016).

Multiple chains methods prove efficient in addressing issues such as local optima, multimodality, and high-dimensional parameter spaces, which are challenges that single chain methods struggle to deal with (Braak, 2006; Vrugt, 2016). Studies by Laloy and Vrugt (2012) and Vrugt (2016) have demonstrated that the DREAM algorithm outperforms the aforementioned adaptive MCMC methods. Consequently, we employ the DREAM algorithm to infer the properties of heterotrophic bacteria and the biodegradability of DOC in the seven water samples collected from the Seine River basin. Detailed descriptions of the aforementioned algorithms, including DREAM, are beyond the scope of this paper, and interested readers are referred to Vrugt (2016) for further information.

2.3.4. Prior distributions of considered parameters and likelihood function used in Bayesian inference

The parameter ranges, or parameter spaces, of the physiological properties and biodegradability of DOC are delineated on the basis of previous sensitivity analyses of the C-RIVE model (Wang et al., 2018; Hasanyar et al., 2023b) and detailed in Table 2. A uniform distribution ($\mathscr{U}_{(min,max)}$) is adopted as the prior distribution (Eq. (10)), $f(\mathbf{x})$) for bacterial physiology (growth rate, growth yield, and mortality rate) and DOC biodegradability while a normal distribution ($f(\mathbf{y}_0) = \mathscr{N}_{(\mathbf{y}_0^*,\sigma_0^2)}$), centered as the measured concentration, is considered for the initial concentrations of DOC and heterotrophic bacteria. σ_0 represents the standard deviation of the observations errors of the initial concentrations. It is fixed at 5 % of the measured values for DOC and at 10 % for bacterial biomass. The parameters and the initial concentrations are assumed to be mutually independent.

In this study, the observation data ($\mathbf{y}^*$) encompass concentrations of DOC ($\mathbf{y}_{doc}^*$) and biomass of heterotrophic bacteria ($\mathbf{y}_{hb}^*$). To address the measurement errors of DOC concentration ($\eta_{t,doc}$) and bacterial biomass ($\eta_{t,hb}$) independently, two coefficients (s_{hb} and s_{doc}) are introduced.

Table 2
Parameter ranges and prior distribution.

Parameters	Unit	Description	Range	Prior
$\mu_{max20,hb}$	[h^{-1}]	Maximum growth rate at 20 °C	[0.01, 0.15]	$\mathscr{U}_{(0.01,0.15)}$
Y_{hb}	[−]	Growth yield at 20 °C	[0.03, 0.50]	$\mathscr{U}_{(0.03,0.5)}$
$k_{d20,hb}$	[h^{-1}]	Mortality rate at 20 °C	[0.01, 0.08]	$\mathscr{U}_{(0.01,0.08)}$
b_1	[−]	BDOC/DOC ratio	[0.05, 0.54]	$\mathscr{U}_{(0.05,0.54)}$
s_1	[−]	DOC_1/BDOC ratio	[0.05, 0.95]	$\mathscr{U}_{(0.05,0.95)}$
hb_{init}	[mgC L^{-1}]	Initial bacteria biomass		$\mathscr{N}_{(\mathbf{y}_0^*,(0.10\mathbf{y}_0^*)^2)}$
DOC_{init}	[mgC L^{-1}]	Initial concentration of DOC		$\mathscr{N}_{(\mathbf{y}_0^*,(0.05\mathbf{y}_0^*)^2)}$
s_{hb}	[−]	Coefficient to determine standard deviations of observation errors of bacteria biomass	[0.05, 0.30]	$\mathscr{U}_{(0.05,0.30)}$
s_{doc}	[−]	Coefficient to determine standard deviations of observation errors of DOC	[0.05, 0.30]	$\mathscr{U}_{(0.05,0.30)}$

BDOC: biodegradable dissolved organic carbon in 45 days
DOC_1: rapidly biodegradable dissolved organic carbon in 5 days

$$\eta_{t,hb} \sim \mathscr{N}_{\left(0,\sigma_{t,hb}^2\right)} \quad \sigma_{t,hb} = s_{hb} \times y_{t,hb}^* \tag{13}$$

$$\eta_{t,doc} \sim \mathscr{N}_{\left(0,\sigma_{t,doc}^2\right)} \quad \sigma_{t,doc} = s_{doc} \times y_{t,doc}^* \tag{14}$$

The likelihood $f(\mathbf{y}_t^*|\mathbf{y}_{1:T})$ can be evaluated using the probability density function of the multivariate normal distribution:

$$f(\mathbf{y}_t^*|\mathbf{y}_{1:T}) \propto \prod_{i\in\{doc,hb\}} \frac{1}{\sigma_i^T} exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{\mathbf{y}_{i,t}^* - \mathbf{y}_{i,t}}{\sigma_i}\right)^2\right). \tag{15}$$

By incorporating the two coefficients s_{hb} and s_{doc} as unknown in the MCMC algorithm, the measurement errors of DOC concentration and bacterial biomass can be inferred. A uniform distribution is adopted as the prior distribution ($s_{hb} \sim \mathscr{U}_{(0.05,0.30)}$ and $s_{doc} \sim \mathscr{U}_{(0.05,0.30)}$). In total, 9 parameters (unknown) are integrated into the DREAM algorithm (Table 2).

2.3.5. Convergence diagnostic of MCMC and model performance evaluation

To evaluate whether the multiple chains have converged to the target distribution, the *potential* scale reduction factor ($\widehat{R}$), proposed by Gelman and Rubin (1992), is computed. This diagnostic involves comparing the between-chain and within-chain variances of each parameter.

$$W = \frac{1}{N(n-1)}\sum_{j=1}^{N}\sum_{i=1}^{n}\left(\mathbf{x}_j^i - \overline{\mathbf{x}}_j\right)^2 \tag{16}$$

$$\frac{B}{n} = \frac{1}{N-1}\sum_{j=1}^{N}\left(\overline{\mathbf{x}}_j - \overline{\mathbf{x}}\right)^2 \tag{17}$$

$$\widehat{V} = \frac{n-1}{n}W + \frac{N+1}{Nn}B \tag{18}$$

$$\widehat{R} = \sqrt{\frac{\widehat{V}}{W}} \tag{19}$$

With N: number of Markov chains.
n: number of iterations of each chain.

W: averaged variances of N chains, within-chain variance.

$\frac{B}{n}$: variance of the means of multiple chains, between-chain variance.

In a state of convergence, both $\widehat{V}$ and W serve as unbiased estimates of the true variance of the target distribution. The original proposal by Gelman and Rubin (1992) suggests running the chains for $2n$ iterations and basing the calculation of $\widehat{R}$ on the final n iterations for each parameter of interest. If $\widehat{R}$ is less than 1.2 for each parameter, then the latter half-sequences have reached approximately convergence to the target distribution (Brooks and Gelman, 1998; Vrugt, 2016). In this paper, the more stringent criteria of $\widehat{R} \leq 1.05$ is adopted.

The RMSE (root mean square error) and pbias (percent bias) are used to evaluate the performance of the HSB model.

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{m} (sim_i - obs_i)^2 / m} \tag{20}$$

$$\mathrm{pbias} = 100 \times \frac{\sum_{i=1}^{m} (sim_i - obs_i)}{\sum_{i=1}^{m} obs_i} \tag{21}$$

With m: number of observations.
sim: simulated concentrations.
obs: observed concentrations.

3. Results

3.1. MCMC setting and convergence diagnostic

A total of 9 parameters (Table 2) require characterization using the DREAM algorithm. Recent research has suggested that the number of chains (N) should be at least twice the dimension of the parameters (d, $N \geq 2d$) for differential evolution-based samplers (Brunetti et al., 2023). In this study, 20 Markov chains, each comprising 50,000 iterations, have been employed to ensure convergence of the DREAM sampler for each water sample. The *potential* scale reduction factors ($\widehat{R}$) indicate that the DREAM samplers achieve convergence rapidly and that 50,000 iterations are deemed sufficient. Convergence was achieved in less than 10,000 iterations (Fig. SI-1).

3.2. Simulated concentrations of DOC and heterotrophic bacteria

After convergence, the last half-sequences of each Markov chain (25,000 iterations) are used to estimate the posterior distributions of the parameters through Kernel Density Estimation. This enables the estimation of the 95 % credibility intervals of the posterior distributions (Fig. 4 and Table SI-1). The optimal parameter values are considered as the values corresponding to the maximum a posteriori (MAP) during these 25000×20 iterations (Table SI-1). Subsequently, the

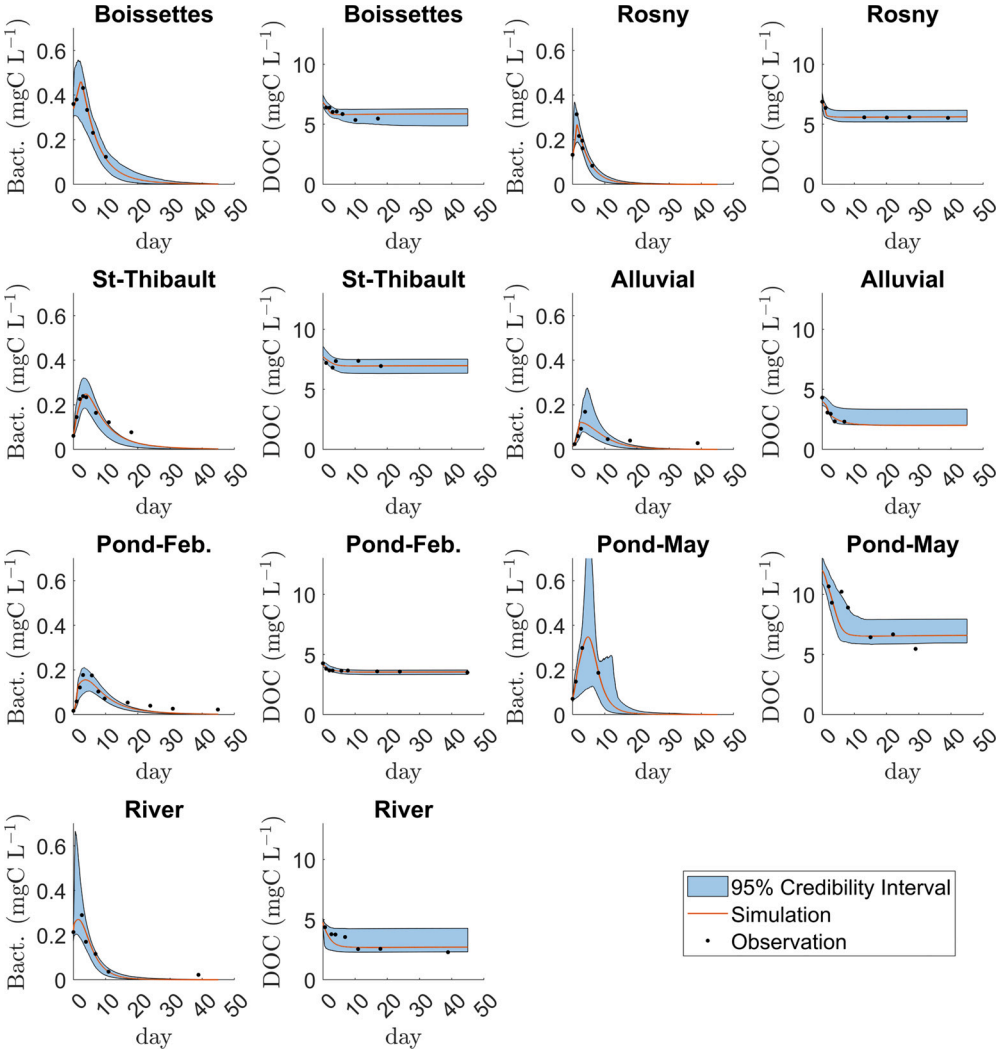


Fig. 2. Simulated concentrations of DOC and heterotrophic bacteria using the optimal parameter values corresponding to the maximum a posteriori (noted as "Simulation" in the graphs) and their 95 % credibility intervals (for 500,000 realizations).

concentrations of DOC and heterotrophic bacteria are simulated using the HSB model and the optimal parameter values (MAP).

The findings demonstrate that the dynamics of DOC biodegradation are effectively simulated with the optimal parameter values (Fig. 2). Heterotrophic bacteria utilize biodegradable DOC for both growth and respiration. Rapid bacterial growth primarily occurs within the initial 5 days, coinciding with the degradation of fast biodegradable DOC (DOC_1, Eqs. (1), (2)). Following this, bacterial growth is constrained by the availability of small monomeric substrate (SMS), which is in turn limited by the hydrolysis of slowly biodegradable DOC (DOC_2, Eqs. (2) and (7)).

The maximum RMSE of simulated DOC concentrations is 1.473 mgC L^{-1} for Pond-May sample. The percent bias of −6.91 % indicates that the simulated concentrations are averagely smaller than observed ones (Fig. 2). The maximum pbias of simulated DOC concentrations is −6.94 % for River sample while its RMSE is calculated at 0.503 mgC L^{-1}. For other samples, the RMSE of simulated DOC concentrations are less than 0.5 mgC L^{-1} and the pbias are not greater than 1 %. The bacterial biomass tend to be underestimated compared to the observations (Table 3). Although relative important pbias are estimated for Alluvial and Pond-Feb. samples (−12.37 % and − 9.85 %), the RMSE of simulated bacterial biomass are around 0.02 mgC L^{-1} for all samples (Table 3). The statistical criteria confirm the high performances of the HSB model with the optimal parameter values (MAP) which characterize the biodegradability of DOC and physiological properties of heterotrophic bacteria.

3.3. Biodegradability of DOC

3.3.1. Biodegradable dissolved organic carbon (BDOC)

The results reveal that the water collected from the hydrosystem (surface water and groundwater) generally contains higher concentrations of BDOC compared to water sampled from the discharge of WWTPs, both in terms of absolute concentrations and as a percentage of DOC (Fig. 3). Specifically, the estimated BDOC concentrations in natural water range from 0.61 mgC L^{-1} (Pond-Feb.) to 5.74 mgC L^{-1} (Pond-May), whereas for treated water, they vary from 0.82 mgC L^{-1} (St-Thibault) to 1.40 mgC L^{-1} (Boissettes). Although the BDOC fraction constitutes averagely 15.7 % of the DOC in treated water, this proportion is lower for the St-Thibault WWTP (10.7 %). Conversely, the biodegradable fraction of DOC (b_1) in hydrosystem water samples is approximately 50 %, except for the water sampled from Pond in February 2021 (15 %, with a 95 % credibility interval of [13.2 %, 20.3 %], Fig. 3 and Table SI-1). Notably, the BDOC fraction is estimated at 47.9 % for the water sampled from Pond in May 2021 (Fig. 3), with a 95 % credibility interval of [33.2 %, 52.6 %] (Table SI-1). The highest percentage of BDOC (represented by the value of b_1) is observed in water sampled from alluvial sources, with a value of 53.8 % (Fig. 3).

3.3.2. Fast biodegradable dissolved organic carbon (BDOC fast)

Concerning the fraction of fast biodegradable DOC within 5 days (BDOC fast, denoted as $s_1 = \frac{DOC_1}{BDOC}$, Eq. (1)), it exhibits considerable variability (Fig. 3). No significant difference can be observed between treated and natural water samples. The estimated fast BDOC (s_1, Eq. (1)) ranges from 9.3 % to 75.2 % of BDOC for water samples collected from WWTPs, while it varies between 6.7 % and 38.4 % of BDOC for hydrosystem water samples. However, the 95 % credibility intervals reveal substantial uncertainties regarding the percentage of BDOC fast (Fig. 4). For instance, the 95 % credibility interval of the fraction of fast BDOC for the River water sample suggests a range from 7.7 % to 86.2 % (Table SI-1 and Fig. 3), which is similar for Boissettes ([6.7 %, 85.4 %]), Rosny ([8.8 %, 84.8 %]) Alluvial ([9.4 %, 83.3 %]), and Pond-May ([5.5 %, 75.4 %]).

3.4. Physiological properties of heterotrophic bacteria

Utilizing the MCMC algorithm, it becomes feasible to characterize the physiological properties of heterotrophic bacteria (i.e., maximal growth rate at 20 °C, growth yield, and mortality rate at 20 °C). The posterior distributions of the mortality rate at 20 °C ($k_{d20,hb}$) indicate predominantly low probable values and exhibit minimal uncertainties across all water samples (Fig. 4). The average of the optimal parameter values for the mortality rate at 20 °C stands at 0.013 h^{-1}.

The estimated growth yield of heterotrophic bacteria (Y_{hb}) displays variability in both treated and natural water samples, ranging from 0.125 to 0.498 for natural water and from 0.179 to 0.491 for treated water respectively (Table SI-1 and Fig. 4). The averaged optimal growth yield of bacteria in treated water appears to surpass that in natural water (0.337 vs. 0.247). Notably, Pond-Feb and St-Thibault WWTPs exhibit the highest optimal values of growth yield (Fig. 4). The posterior distributions of the maximum growth rate at 20 °C ($\mu_{max20,hb}$) for Rosny, St-Thibault, and River depict wide 95 % credibility intervals, indicative of substantial uncertainties (Fig. 4). The average optimal value of $\mu_{max20,hb}$ is determined as 0.061 h^{-1}.

4. Discussions

4.1. Uncertainties of parameterization of DOC biodegradability in the HSB model

In the HSB model, DOC is partitioned into three pools based on their biodegradability rates (Section 2.2.2). The results reveal that the parameter governing the fraction of fast biodegradable DOC within 5 days (s_1) exhibits greater uncertainties (Fig. 4) compared to the fraction of biodegradable DOC (b_1). The posterior distributions of s_1 are notably dispersed, displaying large 95 % credibility intervals (Fig. 4). Furthermore, a wide posterior distribution, such as a uniform distribution, suggests minimal influence of the parameter on the simulation outcomes. A recent sensitivity analysis highlighted that the fraction of biodegradable DOC (b_1) exerts more significant influence than the fraction of fast biodegradable DOC (s_1) on dissolved oxygen levels during low flow conditions (Hasanyar et al., 2023b). Particularly, the fraction of biodegradable DOC (b_1) emerges as the second most influential parameter under high bacteria net growth conditions (Hasanyar et al., 2023b). The marginal posterior distributions obtained in this study through experiments reaffirm the findings of the sensitivity analysis conducted by Hasanyar et al. (2023b) on oxygen data from river water.

In river water quality modeling, determining the seven fractions of total organic carbon (Section 2.2.2) for incoming inflows like river tributaries, combined sewer overflow, and WWTP discharge is often challenging. To mitigate uncertainties regarding the fraction of fast biodegradable DOC (s_1), a variant of the HSB model with five organic carbon pools instead of seven could be alternatively used (Wang et al., 2020). This approach involves grouping slowly and rapidly biodegradable organic carbon ($BDOC = DOC_1 + DOC_2$ and $BPOC = POC_1 +$

Table 3
Root mean square error (RMSE, mgC L^{-1}) and percent bias (pbias, %) for concentrations of DOC and bacteria.

Criteria	Boissettes	Rosny	St-Thibault	Alluvial	Pond-Feb.	Pond-May	River
$RMSE_{doc}$	0.285	0.109	0.312	0.258	0.099	1.473	0.503
$pbias_{doc}$	0.26	−0.14	−0.16	0.65	−0.20	−6.91	−6.94
$RMSE_{hb}$	0.013	0.024	0.020	0.027	0.021	0.006	0.024
$pbias_{hb}$	0.44	−3.13	−3.43	−12.37	−9.85	−0.06	−2.69

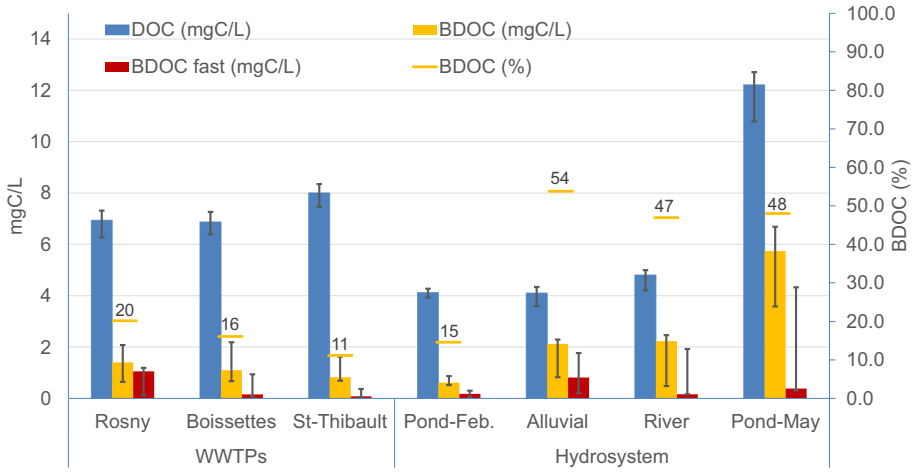


Fig. 3. Synthesis of the properties of DOC biodegradability of the seven water samples. The black bars represent 95 % credibility intervals.

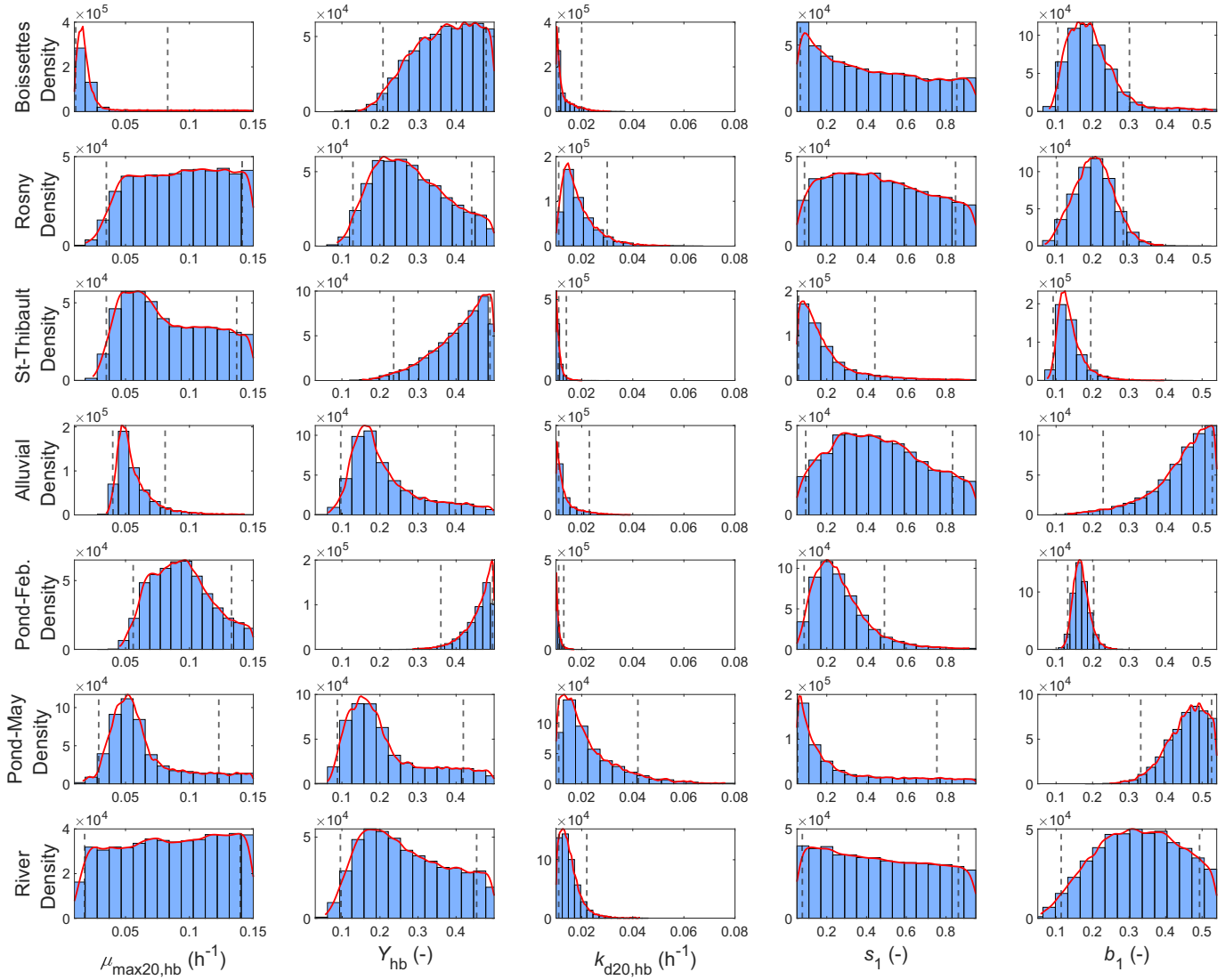


Fig. 4. Marginal posterior distributions of the maximal growth rate at 20 °C ($\mu_{max20,hb}$), the growth yield (Y_{hb}), the mortality rate at 20 °C ($k_{d20,hb}$), the fraction of fast biodegradable in BDOC (s_1), and the fraction of biodegradable DOC (b_1). Black dashed lines indicate the 95 % credibility intervals.

POC_2). The five organic carbon pools in this setup include the small monomeric substrate (SMS), the biodegradable DOC (BDOC = $b_1 \times$ DOC), the biodegradable POC (BPOC), and the refractory organic carbon (RDOC = $(1 - b_1) \times$ DOC and RPOC). However, this adjustment may affect other parameter values, such as the half-saturation constant for small monomeric substrate ($K_{sms,hb}$, Eq. (4)), the maximum hydrolysis rates of biodegradable organic carbon, and potentially the physiological properties of heterotrophic bacteria (Wang et al., 2020). Therefore, further experiments and investigations would be necessary in future studies to confirm the relevance of physiological parameters and the robustness of the HSB model with five vs. seven organic carbon pools.

4.2. Uncertainties of physiological properties of heterotrophic bacteria in the HSB model

4.2.1. Averaged physiological properties of heterotrophic bacteria

The posterior distributions of the mortality rate at 20 °C ($k_{d20,hb}$) exhibit pronounced peaks, compared to growth rate and growth yield (Fig. 4). This observation aligns with the recent sensitivity analysis of the C-RIVE model, which identified the mortality rate at 20 °C as the most influential parameter affecting dissolved oxygen concentration under low-flow conditions (Hasanyar et al., 2023b). The average optimal value of $k_{d20,hb}$ (0.013 h^{-1}) closely resembles the disappearance rate of small heterotrophic bacteria documented by Garnier et al. (1992b) in the Seine River (0.0125 h^{-1}). Likewise, the average of the maximum growth rates at 20 °C of 0.061 h^{-1} ($\mu_{max20,hb}$) aligns with the value of 0.065 h^{-1} reported by Garnier et al. (1992b) for the Seine River. This consistency suggests the dominance of small heterotrophic bacteria (Garnier et al., 1992a) in the seven water samples, corroborated by results from epifluorescence microscopy. In addition, the averaged optimal growth yield (Y_{hb}) is estimated at 0.34 for treated water and at 0.25 for natural water, which is coherent with the historical reference growth yield value of 0.33 ± 0.06 experimentally determined (Bariller and Garnier, 1993) and 0.25 mostly used in the HSB model for natural waters (Billen et al., 1994; Wang et al., 2024).

4.2.2. Parameter correlation: Growth yield (Y_{hb}), mortality rate at 20 °C ($k_{d20,hb}$), and biodegradability of DOC (b_1)

A scatter plot of all parameters enables the investigation of parameter correlations. The results reveal positive correlation (Fig. SI-2) between the growth yield (Y_{hb}) and the mortality rate at 20 °C ($k_{d20,hb}$) for the Rosny, Alluvial, and Pond-May water samples with correlation coefficients of 0.63, 0.72, and 0.83 respectively. Conversely, a negative correlation between the growth yield (Y_{hb}) and biodegradability of DOC (b_1) has been observed for the St-Thibault, Pond-Feb, and River samples (correlation coefficients of −0.82, −0.55, and − 0.58 respectively).

These scatter plots show strong nonlinear interactions between parameters (banana shaped, Fig. SI-2). The substantial interactions of the growth yield (Y_{hb}), mortality rate at 20 °C ($k_{d20,hb}$), and biodegradability of DOC (b_1) with other parameters have been illustrated by Sobol sensitivity analyses (Wang et al., 2018; Hasanyar et al., 2023b). These interactions present challenges for calibrating water quality models or applying data assimilation techniques (Wang et al., 2019, 2022) due to the concept of equifinality (Beven, 1989, 2006). The posterior distributions of $k_{d20,hb}$ show pronounced peaks with similar optimal values for all water samples. Therefore, the authors recommend fixing the mortality rate at 20 °C of the small heterotrophic bacteria (Garnier et al., 1992b) to a constant value of 0.013 h^{-1} for future data assimilation applications.

4.3. Uncertainties of the measurement of DOC concentrations and bacterial biomass

4.3.1. Initial DOC concentrations and initial bacterial biomass

The marginal posterior distributions of initial bacterial biomass (hb_{init}) are consistent with the prior distributions (Fig. SI-3), which indicates that the prior distributions of initial bacterial biomass are well defined. The standard deviation of observation errors of the initial bacterial biomass was set to 10 % of the measured values (Table 2), which can then be considered as reasonable. Generally, the prior distributions of the initial DOC concentrations are well defined also. Specifically, the measured initial DOC concentrations for Alluvial and Pond-Feb. samples appear to be higher than the posterior values. The MCMC method enables thus the inference of the initial DOC concentration using water incubation data and refinement of the prior distributions of DOC_{init} (Fig. SI-3).

4.3.2. Measurements during water incubation

The two parameters (s_{hb} and s_{doc}, Table 2) to determine the standard deviation of observation errors allow to originally explore the measurement and model uncertainties during water incubation, which were rarely quantified in water quality modeling. The results show that the standard deviation of measurement uncertainties of bacterial biomass represents averagely 19.5 % of the measured values (Table SI-2), which corresponds to 0.0245 mgC L^{-1}. For DOC, standard deviation of measurement uncertainties accounts for 9 % of the measured concentrations, and a concentration of 0.493 mgC/L.

4.4. Recommendations on parameter ranges for future studies

Parameter ranges play a crucial role in sensitivity analysis and data assimilation simulations. A recent study highlights the potential for quantifying DOC biodegradability (b_1, Table 2) through oxygen data assimilation (Hasanyar et al., 2023a, preprint). However, employing a shared DOC biodegradability parameter range for different water inflows (river tributaries, combined sewer overflow, WWTP discharge) lacks realism, as variations in DOC composition and sources significantly impact biodegradability (Liu et al., 2021; Liu and Wang, 2022; Begum et al., 2023) and microbial processes (McCarren et al., 2010; Muscarella et al., 2019; Zhou et al., 2024).

This study proposes distinct parameter ranges for DOC biodegradability and the physiological properties of heterotrophic bacteria based on water origins (treated water - WWTPs or natural water - Hydrosystem). For each water type (WWTP or Hydrosystem), the parameter ranges (Table 4) are determined based on the minimum and maximum levels of the 95 % credibility intervals of posterior distributions (Table SI-1). These parameter ranges can be employed in future data assimilation efforts to differentiate the DOC biodegradability of various sources.

Certainly, the parameter ranges presented in this section could benefit from further refinement through the inclusion of additional incubation data from a broader spectrum of water sources and seasons. For instance, the incorporation of combined sewer overflow data, which remains unexplored, would enhance the simulation of river metabolism during extreme events such as storms (Wang et al., 2022).

4.5. Benefit of bacterial information during water incubation experiment for biogeochemical modeling

In practice, biogeochemists conducting water incubation

Table 4
Estimated parameter ranges related to the physiological properties of heterotrophic bacteria and DOC biodegradability for treated water (WWTPs) and natural water (Hydrosystem).

Parameters	$\mu_{max20,hb}$	Y_{hb}	b_1
	[h^{-1}]	[−]	[−]
WWTPs	[0.011, 0.141]	[0.129, 0.489]	[0.092, 0.301]
Hydrosystem	[0.018, 0.140]	[0.088, 0.496]	[0.114, 0.528]

experiments typically focus on the biodegradable fraction of DOC (Liu et al., 2021; Liu and Wang, 2022; Begum et al., 2023) rather than directly measuring bacterial biomass. Consequently, until now the MCMC method has been exclusively applied with only DOC data to explore the uncertainties in predicting the biomass and physiological properties of heterotrophic bacteria. The results demonstrate that the DOC concentrations can be accurately simulated (not shown here). However, the uncertainties associated the bacterial biomass are notably high due to the absence of observed bacterial information in the MCMC algorithm. In this context, while the biodegradable fraction of DOC can be effectively determined, the physiological properties of heterotrophic bacteria remain highly uncertain. These findings underscore the importance of incorporating bacterial biomass information into MCMC analyses to infer the physiological properties of heterotrophic bacteria (Garnier et al., 1992a), which are essential for a comprehensive biogeochemical modeling (Wang et al., 2024).

5. Conclusions

The paper employs a multiple Markov Chain Monte Carlo method (DREAM) to infer the physiological properties of heterotrophic bacteria and the biodegradability of dissolved organic carbon in the Seine River basin, France. To achieve this, seven water samples were collected from either the hydrosystem (river/standing water; groundwater - alluvial) or from the discharge of wastewater treatment plants in 2021. These water samples were inoculated and incubated at an ambient temperature of 20 °C. The concentrations of DOC and heterotrophic bacteria were monitored over a 45-day period. The HSB model was utilized to simulate the dynamics of organic carbon degradation during the incubation period. Based on the results, the uncertainties of bacterial physiology and DOC biodegradability, the parameter correlations, the measurement uncertainties, and the parameter ranges are discussed. The seven contrasted water samples used in the paper serve as a proof of concept. They indeed display the potential of the method to identify parameter values in a large range of value. We propose a first synthesis of those values in the form of narrowed parameter ranges by sample origin. This methodology can be extended to analyze additional water incubation datasets in future research by experimentalists focused on the carbon cycle in aquatic environment or applied to previous samples with existing data, enabling a deeper understanding of the influencing factors of DOC/BDOC and the uncertainties in model parameters.

The following conclusions emerge from the study.

- The combination of laboratory experiments (water incubation) and modeling (HSB model) using a Bayesian approach is effective for investigating the physiological properties of heterotrophic bacteria, DOC biodegradability, and the uncertainties in parameters and measurements.
- The biodegradable fraction of DOC exhibits a notably higher presence in the natural water samples compared to the treated water samples. Additionally, the fraction of fast biodegradable DOC within a 5-day period displays considerable variability across both water sample types. The wide posterior distributions underscore substantial uncertainties/challenges in parameterizing the fraction of fast biodegradable DOC within the HSB model. As a potential solution, a version of the HSB model with five pools of organic carbon, rather than the original seven, is proposed.
- The determined mortality rate of heterotrophic bacteria at 20 °C exhibits remarkable stability at 0.013 h^{-1}, with no statistically significant differences observed between natural and treated water samples. Regarding the parameter correlations, the authors recommend fixing the mortality rate at 20°C to 0.013 h^{-1} in future data assimilation applications.
- Treated water is characterized by higher growth yields of heterotrophic bacteria compared to natural water (0.34 vs. 0.25).
- Bacterial biomass data is necessary for inferring the physiological properties of heterotrophic bacteria using the MCMC method.

CRediT authorship contribution statement

Shuaitao Wang: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **Nicolas Flipo:** Writing – review & editing, Validation, Methodology, Conceptualization. **Josette Garnier:** Writing – review & editing, Methodology, Investigation. **Thomas Romary:** Writing – review & editing, Validation, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was carried out under the PIREN-Seine program (https://www.piren-seine.fr/), part of the french Long Term Socio-Ecological Research (LTSER) (named in France "Zone Atelier Seine" (INEE, CNRS). We would like to express our gratitude to Lou Weidenfeld, Benjamin Mercier, and Anun Marninez for conducting the water incubation experiments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.177252.

Data availability

Data will be made available on request.

References

Abbott, B.W., Larouche, J.R., Jones Jr., J.B., Bowden, W.B., Balser, A.W., 2014. Elevated dissolved organic carbon biodegradability from thawing and collapsing permafrost. J. Geophys. Res. Biogeosci. 119, 2049–2063. https://doi.org/10.1002/2014JG002678.

Bariller, A., Garnier, J., 1993. Influence of temperature and substrate concentration on bacterial growth yield in seine river water batch cultures. Appl. Environ. Microbiol. 59, 1678–1682.

Battin, T.J., Luyssaert, S., Kaplan, L.A., Aufdenkampe, A.K., Richter, A., Tranvik, L.J., 2009. The boundless carbon cycle. Nat. Geosci. 2, 598–600.

Battin, T.J., Lauerwald, R., Bernhardt, E.S., Bertuzzo, E., Gener, L.G., Hall, R.O., Hotchkiss, E.R., Maavara, T., Pavelsky, T.M., Ran, L., Raymond, P., Rosentreter, J.A., Regnier, P., 2023. River ecosystem metabolism and carbon biogeochemistry in a changing world. Nature 613, 449–459. https://doi.org/10.1038/s41586-022-05500-8.

Bayes, T., 1763. An essay towards solving a problem in the doctrine of chances. Phil. Trans. Royal Soc. London 53, 370–418.

Begum, M.S., Jang, I., Lee, J.M., Oh, H.B., Jin, H., Park, J.H., 2019. Synergistic effects of urban tributary mixing on dissolved organic matter biodegradation in an impounded river system. Sci. Total Environ. 676, 105–119. https://doi.org/10.1016/j.scitotenv.2019.04.123.

Begum, M.S., Park, J.H., Yang, L., Shin, K.H., Hur, J., 2023. Optical and molecular indices of dissolved organic matter for estimating biodegradability and resulting carbon dioxide production in inland waters: a review. Water Res. 228, 119362. https://doi.org/10.1016/j.watres.2022.119362.

Beven, K., 1989. Changing ideas in hydrology. The case of physically-based model. J. Hydrol. 105, 157–172.

Beven, K., 2006. A manifesto for the equifinality thesis. J. Hydrol. 320, 18–36.

Billen, G., 1991. Protein degradation in aquatic environments. In: Chrost, R. (Ed.), Microbial Enzyme in Aquatic Environments. Springer Verlag, Berlin.

Billen, G., Servais, P., 1989. Modélisation des processus de dégradation bactérienne de la matière organique en milieu aquatique, in: et al., B. (Ed.), Micro-organismes dans les écosystèmes océaniques. Masson Paris, pp. 219–245.

Billen, G., Garnier, J., Hanset, P., 1994. Modelling phytoplankton development in whole drainage networks: the RIVERSTRAHLER model applied to the seine river system. Hydrobiologia 289, 119–137.

Bottero, A., Gesret, A., Romary, T., Noble, M., Maisons, C., 2016. Stochastic seismic tomography by interacting markov chains. Geophys. J. Int. 207, 374–392.
Braak, C.J.F.T., 2006. A markov chain Monte Carlo version of the genetic algorithm differential evolution: easy bayesian computing for real parameter spaces. Stat. Comput. 16, 239–249. https://doi.org/10.1007/s11222-006-8769-1.
Brooks, S., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. 7, 434–455.
Brunetti, G., Šimunek, J., Wöhling, T., Stumpp, C., 2023. An in-depth analysis of markov-chain Monte Carlo ensemble samplers for inverse vadose zone modeling. J. Hydrol. 624, 129822. https://doi.org/10.1016/j.jhydrol.2023.129822.
Butman, D., Raymond, P.A., 2011. Significant efflux of carbon dioxide from streams and rivers in the United States. Nat. Geosci. 4, 839–842. https://doi.org/10.1038/ngeo1294.
Cole, J., Prairie, Y., Caraco, N., Mc Dowell, W., Tranvik, L., Striefl, R., Duarte, C., Kortelainen, P., Downing, J., Middelburg, J., Melack, J., 2007. Plumbing the global carbon cycle: Intergrating inland waters into the terrestrial carbon budget. Ecosystems 10, 171–184.
Deemer, B.R., Harrison, J.A., Li, S., Beaulieu, J.J., DelSontro, T., Barros, N., Bezerra-Neto, J.F., Powers, S.M., dos Santos, M.A., Vonk, J.A., 2016. Greenhouse gas emissions from reservoir water surfaces: a new global synthesis. BioScience 66, 949–964. https://doi.org/10.1093/biosci/biw117.
Drake, T.W., Raymond, P.A., Spencer, R.G.M., 2018. Terrestrial carbon inputs to inland waters: a current synthesis of estimates and uncertainty. Limnol. Oceanogr. Letters 3, 132–142. https://doi.org/10.1002/lol2.10055.
Even, S., Poulin, M., Garnier, J., Billen, G., Servais, P., Chesterikoff, A., Coste, M., 1998. River ecosystem modelling: application of the ProSe model to the seine river (France). Hydrobiologia 373, 27–45. https://doi.org/10.1023/A:1017045522336.
Even, S., Poulin, M., Mouchel, J.M., Seidl, M., Servais, P., 2004. Modelling oxygen deficits in the seine river downstream of combined sewer overflows. Ecol. Model. 173, 177–196.
Even, S., Mouchel, J.M., Servais, P., Flipo, N., Poulin, M., Blanc, S., Chabanel, M., Paffoni, C., 2007. Modeling the impacts of combined sewer overflows on the river seine water quality. Sci. Total Environ. 375, 140–151. https://doi.org/10.1016/j.scitotenv.2006.12.007.
Fellman, J., Hood, E., Spencer, R., 2010. Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: a review. Limnol. Oceanogr. 55, 2452–2462.
Fellman, J.B., D'Amore, D.V., Hood, E., Boone, R.D., 2008. Fluorescence characteristics and biodegradability of dissolved organic matter in forest and wetland soils from coastal temperate watersheds in Southeast Alaska. Biogeochemistry 88, 169–184. https://doi.org/10.1007/s10533-008-9203-x.
Flipo, N., Even, S., Poulin, M., Tusseau-Vuillemin, M.H., Améziane, T., Dauta, A., 2004. Biogeochemical modelling at the river scale: plankton and periphyton dynamics - grand Morin case study. France. Ecol. Model. 176, 333–347.
Flipo, N., Rabouille, C., Poulin, M., Even, S., Tusseau-Vuillemin, M.H., Lalande, M., 2007. Primary production in headwater streams of the seine basin: the grand Morin case study. Sci. Total Environ. 375, 98–109. https://doi.org/10.1016/j.scitotenv.2006.12.015.
Garnier, J., Billen, G., 1994. Ecological interactions in a shallow sand-pit lake (Lake créteil, parisian basin, France): a modelling approach. Hydrobiologia 275 (276), 97–114.
Garnier, J., Billen, G., Servais, P., 1992a. Physiological characteristics and ecological role of small- and large-sized bacteria in a polluted river (seine river, France). Arch. Hydrobiol. Beih. 37, 83–94.
Garnier, J., Servais, P., Billen, G., 1992b. Bacterioplankton in the seine river (France): impact of the Parisian urban effluent. Can. J. Microbiol. 38, 56–64.
Garnier, J., Billen, G., Coste, M., 1995. Seasonal succession of diatoms and chlorophycae in the drainage network of the river seine: observations and modelling. Limnol. Oceanogr. 40, 750–765.
Garnier, J., Billen, G., Sanchez, N., Leporcq, B., 2000. Ecological functioning of the Marne reservoir (upper seine basin, France). Regul. Rivers: Res. Mgmt. 16, 51–71.
Garnier, J., Ramarson, A., Billen, G., Théry, S., Thiéry, D., Thieu, V., Minaudo, C., Moatar, F., 2018. Nutrient inputs and hydrology together determine biogeochemical status of the loire river (France): current situation and possible future scenarios. Sci. Total Environ. 637-638, 609–624. https://doi.org/10.1016/j.scitotenv.2018.05.045.
Garnier, J., Weidenfeld, L., Billen, G., Martinez, A., Mercier, B., Rocher, V., Tabuchi, J.P., Azimi, S., 2021. La matière organique dans le continuum terrestre-aquatique du bassin de la Seine. PIREN-Seine, Rapport annuel PIREN-Seine. https://doi.org/10.26047/PIREN.rapp.ann.2021.vol21.
Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7, 457–472. http://www.jstor.org/stable/2246093.
Goffin, A., Guérin, S., Rocher, V., Varrault, G., 2017. Caractérisation de l'évolution de la matière organique dissoute de l'amont à l'aval de l'agglomération parisienne pendant une année hydrologique par spectrométrie de fluorescence 3D. Technical Report, PIREN Seine.
Haario, H., Saksman, E., Tamminen, J., 1999. Adaptive proposal distribution for random walk metropolis algorithm. Comput. Stat. 14, 375–395. https://doi.org/10.1007/s001800050022.
Haario, H., Laine, M., Mira, A., Saksman, E., 2006. Dram : efficient adaptive MCMC. Stat. Comput. 16, 339–354.
Hao, X., Ruihong, Y., Zhuangzhuang, Z., Zhen, Q., Xixi, L., Tingxi, L., Ruizhong, G., 2021. Greenhouse gas emissions from the water–air interface of a grassland river: a case study of the xilin river. Sci. Rep. 11, 2659. https://doi.org/10.1038/s41598-021-81658-x.
Hasanyar, M., Flipo, N., Romary, T., WANG, S., 2023a. Quantifying heterotrophic bacteria parameters and dissolved organic carbon biodegradability through oxygen data assimilation in a river water quality model doi: 10.22541/essoar.168987141.18902487/v1.
Hasanyar, M., Romary, T., Wang, S., Flipo, N., 2023b. How much do bacterial growth properties and biodegradable dissolved organic matter control water quality at low flow? Biogeosciences 20, 1621–1633. https://doi.org/10.5194/bg-20-1621-2023.
Hastings, W.K., 1970. Monte Carlo sampling methods using markov chains and their applications. Biometrika 57, 97–109.
Holmes, R.M., McClelland, J.W., Raymond, P.A., Frazer, B.B., Peterson, B.J., Stieglitz, M., 2008. Lability of doc transported by alaskan rivers to the arctic ocean. Geophys. Res. Lett. 35. https://doi.org/10.1029/2007GL032837.
Hotchkiss, E.R., Hall Jr., R.O., Sponseller, R.A., Butman, D., Klaminder, J., Laudon, H., Rosvall, M., Karlsson, J., 2015. Sources of and processes controlling co2 emissions change with the size of streams and rivers. Nat. Geosci. 8, 696–699. https://doi.org/10.1038/ngeo2507.
Laloy, E., Vrugt, J.A., 2012. High-dimensional posterior exploration of hydrologic models using multiple-try dream(zs) and high-performance computing. Water Resour. Res. 48. https://doi.org/10.1029/2011WR010608.
Larouche, J.R., Abbott, B.W., Bowden, W.B., Jones, J.B., 2015. The role of watershed characteristics, permafrost thaw, and wildfire on dissolved organic carbon biodegradability and water chemistry in arctic headwater streams. Biogeosciences 12, 4221–4233. https://doi.org/10.5194/bg-12-4221-2015.
Le, T.P.Q., Billen, G., Garnier, J., Chau, V., 2015. Long-term biogeochemical functioning of the Red River (Vietnam): past and present situations. Reg. Environ. Chang. 15, 329–339. https://doi.org/10.1007/s10113-014-0646-4.
Li, P., Hur, J., 2017. Utilization of uv-Vis spectroscopy and related data analyses for dissolved organic matter (dom) studies: a review. Crit. Rev. Environ. Sci. Technol. 47, 131–154. https://doi.org/10.1080/10643389.2017.1309186.
Liu, F., Wang, D., 2022. Dissolved organic carbon concentration and biodegradability across the global rivers: a meta-analysis. Sci. Total Environ. 818, 151828. https://doi.org/10.1016/j.scitotenv.2021.151828.
Liu, F., Wang, D., Zhang, B., Huang, J., 2021. Concentration and biodegradability of dissolved organic carbon derived from soils: a global perspective. Sci. Total Environ. 754, 142378. https://doi.org/10.1016/j.scitotenv.2020.142378.
Marescaux, A., Thieu, V., Gypens, N., Silvestre, M., Garnier, J., 2020. Modeling inorganic carbon dynamics in the seine river continuum in France. Hydrol. Earth Syst. Sci. 24, 2379–2398. https://doi.org/10.5194/hess-24-2379-2020.
McCarren, J., Becker, J.W., Repeta, D.J., Shi, Y., Young, C.R., Malmstrom, R.R., Chisholm, S.W., DeLong, E.F., 2010. Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. Proc. Natl. Acad. Sci. 107, 16420–16427. https://doi.org/10.1073/pnas.1010732107.
Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092. https://doi.org/10.1063/1.1699114.
Michaelis, L., Menten, M.L., 1913. Die kinetik der invertinwirkung - the kinetics of invertin action. Biochem. Z. 49, 333–369.
Monod, J., 1949. The growth of bacterial cultures. Ann. Rev. Microbiol. 3, 371–394. https://doi.org/10.1146/annurev.mi.03.100149.002103.
Muscarella, M.E., Boot, C.M., Broeckling, C.D., Lennon, J.T., 2019. Resource heterogeneity structures aquatic bacterial communities. ISME J. 13, 2183–2195. https://doi.org/10.1038/s41396-019-0427-7.
Mutschlecner, A.E., Guerard, J.J., Jones, J.B., Harms, T.K., 2018. Regional and intra-annual stability of dissolved organic matter composition and biolability in high-latitude alaskan rivers. Limnol. Oceanogr. 63, 1605–1621. https://doi.org/10.1002/lno.10795.
O'Donnell, J.A., Aiken, G.R., Butler, K.D., Guillemette, F., Podgorski, D.C., Spencer, R.G.M., 2016. Dom composition and transformation in boreal forest soils: the effects of temperature and organic-horizon decomposition state. J. Geophys. Res. Biogeosci. 121, 2727–2744. https://doi.org/10.1002/2016JG003431.
Prairie, Y.T., Alm, J., Beaulieu, J., Barros, N., Battin, T., Cole, J., del Giorgio, P., DelSontro, T., Guérin, F., Harby, A., Harrison, J., Mercier-Blais, S., Serça, D., Sobek, S., Vachon, D., 2018. Greenhouse gas emissions from freshwater reservoirs: what does the atmosphere see? Ecosystems 21, 1058–1071. https://doi.org/10.1007/s10021-017-0198-9.
Price, K., Storn, R.M., Lampinen, J.A., 2005. Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series). Springer-Verlag, Berlin, Heidelberg.
Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., Butman, D., Striegl, R., Mayorga, E., Humborg, C., Kortelainen, P., Dürr, H., Meybeck, M., Ciais, P., Guth, P., 2013. Global carbon dioxide emissions from inland waters. Nature 503, 355–359. https://doi.org/10.1038/nature12760.
Robert, C.P., Casella, G., 2004. Monte-Carlo Statistical Methods, 2nd edition, ed. Springer, Berlin.
Romary, T., 2010. Bayesian inversion by parallel interacting markov chains. Inverse Problems Sci. Eng. 18, 111–130.
Romero, E., Garnier, J., Billen, G., Ramarson, A., Riou, P., Le Gendre, R., 2019. Modeling the biogeochemical functioning of the seine estuary and its coastal zone: export, retention, and transformations. Limnol. Oceanogr. 64, 895–912. doi:https://doi.org/10.1002/lno.11082, arXiv:https://aslopubs.onlinelibrary.wiley.com/doi/pdf/https://doi.org/10.1002/lno.11082.
Servais, P., Billen, G., Vives-Rego, J., 1985. Rate of bacterial mortality in aquatic environments. Appl. Environ. Microbiol. 49, 1448–1454.
Servais, P., Billen, G., Hascoët, M.C., 1987. Determination of the biodegradable fraction of dissolved organic matter in waters. Water Res. 21, 445–450.

Servais, P., Barillier, A., Garnier, J., 1995. Determination of the biodegradable fraction of dissolved and particulate organic carbon in waters. Ann. Limnol. Int. J. Limnol. 31, 75–80. https://doi.org/10.1051/limn/1995005.

Søndergaard, M., Worm, J., 2001. Measurement of biodegradable dissolved organic carbon (bdoc) in lake water with a bioreactor. Water Res. 35, 2505–2513. https://doi.org/10.1016/S0043-1354(00)00532-7.

Stanley, E.H., Casson, N.J., Christel, S.T., Crawford, J.T., Loken, L.C., Oliver, S.K., 2016. The ecology of methane in streams and rivers: patterns, controls, and global significance. Ecol. Monogr. 86, 146–171. https://doi.org/10.1890/15-1027.

Storn, R., Price, K., 1997. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J. Glob. Optim. 11, 341–359. https://doi.org/10.1023/A:1008202821328.

Thieu, V., Billen, G., Garnier, J., 2009. Nutrient transfer in three contrasting NW European watersheds: the seine, Somme, and Scheldt Rivers. A comparative application of the Seneque/Riverstrahler model. Water Res. 43, 1740–1754. https://doi.org/10.1016/j.watres.2009.01.014.

Thieu, V., Mayorga, E., Billen, G., Garnier, J., 2010. Subregional and downscaled global scenarios of nutrient transfer in river basins: seine-somme-scheldt case study. Glob. Biogeochem. Cycles 24. doi:https://doi.org/10.1029/2009GB003561, arXiv: https://agupubs.onlinelibrary.wiley.com/doi/pdf/https://doi.org/10.1029/2009GB003561.

Vilmin, L., Flipo, N., Escoffier, N., Rocher, V., Groleau, A., 2016. Carbon fate in a large temperate human-impacted river system: focus on benthic dynamics. Glob. Biogeochem. Cycles 30, 1086–1104. https://doi.org/10.1002/2015GB005271.

Vonk, J.E., Tank, S.E., Mann, P.J., Spencer, R.G.M., Treat, C.C., Striegl, R.G., Abbott, B. W., Wickland, K.P., 2015. Biodegradability of dissolved organic carbon in permafrost soils and aquatic systems: a meta-analysis. Biogeosciences 12, 6915–6930. https://doi.org/10.5194/bg-12-6915-2015.

Vrugt, J.A., 2016. Markov chain Monte Carlo simulation using the DREAM software package: theory, concepts, and MATLAB implementation. Environ. Model. Softw. 75, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013.

Vrugt, J.A., ter Braak, C.J.F., Gupta, H.V., Robinson, B.A., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch. Env. Res. Risk A. 23, 1011–1026. https://doi.org/10.1007/s00477-008-0274-y.

Wang, S., Flipo, N., Romary, T., 2018. Time-dependent global sensitivity analysis of the C-RIVE biogeochemical model in contrasted hydrological and trophic contexts. Water Res. 144, 341–355. https://doi.org/10.1016/j.watres.2018.07.033.

Wang, S., Flipo, N., Romary, T., 2019. Oxygen data assimilation for estimating micro-organism communities' parameters in river systems. Water Res. 165, 115021. https://doi.org/10.1016/j.watres.2019.115021.

Wang, S., Thieu, V., Flipo, N., Silvestre, M., Weidenfeld, L., Billen, G., 2020. Dégradation de la matière organique dans le modèle de biogéochimie aquatique RIVE : exploration des codes et expérimentation numérique en conditions contrôlées. PIREN Seine, Technical Report. https://doi.org/10.26047/PIREN.rapp.ann.2020.vol02.

Wang, S., Flipo, N., Romary, T., Hasanyar, M., 2022. Particle filter for high frequency oxygen data assimilation in river systems. Environ. Model Softw. 105382. https://doi.org/10.1016/j.envsoft.2022.105382.

Wang, S., Flipo, N., Romary, T., 2023. Which filter for data assimilation in water quality models? Focus on oxygen reaeration and heterotrophic bacteria activity. J. Hydrol. 620, 129423. https://doi.org/10.1016/j.jhydrol.2023.129423.

Wang, S., Thieu, V., Billen, G., Garnier, J., Silvestre, M., Marescaux, A., Yan, X., Flipo, N., 2024. The community-centered freshwater biogeochemistry model unified rive v1.0: a unified version for water column. Geosci. Model Dev. 17, 449–476. https://doi.org/10.5194/gmd-17-449-2024.

Wickland, K.P., Neff, J.C., Aiken, G.R., 2007. Dissolved organic carbon in alaskan boreal forest: sources, chemical characteristics, and biodegradability. Ecosystems 10, 1323–1340. https://doi.org/10.1007/s10021-007-9101-4.

Wickland, K.P., Aiken, G.R., Butler, K., Dornblaser, M.M., Spencer, R.G.M., Striegl, R.G., 2012. Biodegradability of dissolved organic carbon in the Yukon river and its tributaries: seasonality and importance of inorganic nitrogen. Glob. Biogeochem. Cycles 26. https://doi.org/10.1029/2012GB004342.

Yan, X., Garnier, J., Billen, G., Wang, S., Thieu, V., 2022. Unravelling nutrient fate and co2 concentrations in the reservoirs of the seine basin using a modelling approach. Water Res. 225, 119135. https://doi.org/10.1016/j.watres.2022.119135.

Zhou, P., Tian, L., Siddique, M.S., Song, S., Graham, N.J.D., Zhu, Y.G., Yu, W., 2024. Divergent fate and roles of dissolved organic matter from spatially varied grassland soils in China during long-term biogeochemical processes. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.3c08046.