How much do bacterial growth properties and biodegradable dissolved organic matter control water quality at low flow?

Masihullah Hasanyar[1], Thomas Romary[1], Shuaitao Wang[2], and Nicolas Flipo[1]

[1]Mines Paris - PSL, Geosciences and Geoengineering Department, 35 Rue Saint-Honoré, 77300 Fontainebleau, France
[2]Sorbonne Université, CNRS, EPHE, UMR Metis, 75005 Paris, France

Correspondence: Masihullah Hasanyar (masihullah.hasanyar@minesparis.psl.eu)

Abstract. Development of accurate water quality modeling tools is necessary for integrated water quality management of river systems. The existing water quality models can simulate dissolved oxygen (DO) concentration quite well during high flow and phytoplankton blooms in rivers; however, there are discrepancies during the summer low-flow season that are assumed to be due to the uncertainties related to the organic matter contribution of the model boundary conditions. Therefore, we ~~used~~ use the C-RIVE biogeochemical model to evaluate the influence of controlling parameters on DO simulations at low flow. Three Sobol sensitivity analyses (SA) ~~were~~ are carried out based on a coarse model pre-analysis whose ~~target was~~ aim is to develop SA scenarios providing a reduction in the number of model parameters and computation cost as well as hiding inter-parameter interactions. The parameters studied are related to bacterial (e.g., bacterial growth rate), organic matter (OM; repartition and degradation of OM into constituent fractions), and physical factors (e.g., reoxygenation of the river due to navigation and wind), whose variation ranges are selected based on a detailed literature review. Bacterial growth and mortality rates are found to be by far the two most influential parameters, followed by bacterial growth yield. More refined SA results indicate that the biodegradable fraction of dissolved organic matter (BDOM) and the bacterial growth yield are the most influential parameters under conditions of a high net bacterial growth rate (= growth rate – mortality rate), while bacterial growth yield is independently dominant in low net growth situations. Based on the results of this study, proposals are made for in situ measurement of BDOM under ~~a dense and well-equipped~~ an urban area water quality monitoring network that ~~could provide~~ provides high-frequency data. The results also indicate the need for bacterial community monitoring in order to detect potential bacterial community shifts after transient events such as combined sewer overflows and post-infrastructure improvement in treatment plants. Furthermore, we discuss the integration of BDOM in data assimilation software for better estimation of BDOM contribution from boundary conditions, which would result in improved water quality modeling.

1 Introduction

Dissolved oxgyen (DO) has been considered the most important indicator of water quality in surface water resources (Odum, 1956; Escoffier et al., 2018) because it integrates the biological functioning of a system as well as the impact of anthropogenic forcing. It is the main variable used to evaluate river metabolism (Odum, 1956; Staehr et al., 2010; Demars et al., 2015) by comparing the gross primary production (GPP) with ecosystem respiration (ER) and defining whether an ecosystem is

25 autotrophic or heterotrophic based on the net ecosystem production (NEP = GPP-ER) being positive or negative, respectively (Garnier et al., 2020). Maintaining a sufficient level of DO is necessary for the overall health of rivers, not only because of the life dependency of water species (Garvey et al., 2007), but also for preventing smell and taste degradation (Bailey and Ahmadi, 2014).

The situation of rivers during low flow is of particular interest since studies have demonstrated that the river water quality
30 during such flow periods is more vulnerable to degradation due to lower dilution rates. This is particularly the case if the river receives organic matter load from wastewater treatment plants (WWTP) and combined sewage overflows (CSO), thereby leading to heterotrophic conditions (Seidl et al., 1998a; Even et al., 2004; Vilmin et al., 2016; Garnier et al., 2020) in the river, where very low DO levels and high fish mortality can be observed. Therefore, river water quality modeling has been one of the main research interests of water quality managers and researchers ever since the use of the very first water quality model
35 (Streeter and Phelps, 1925) to more complex ones (Even et al., 1998; Flipo, 2005; Wang et al., 2013), aiming to identify the main determinants of DO evolution and to forecast the response of aquatic systems to human-induced pressure, in particular due to releases of ~~the~~ wastewater plants.

Large discrepancies exist between DO simulations and observations during low-flow periods and most models are not able to simulate the evolution of DO accurately. In water quality modeling studies at low flow, the QUESTOR model was applied
40 on the Thames (UK), which demonstrated discrepancies between observed and simulated DO at low flow (Hutchins et al., 2020) and uncertainties related to benthic respiration were revealed to be the main reason for the mismatch. The Riverstrahler model was applied at low flow on the Mosel (Germany) (Garnier et al., 1999), the Scheldt (Belgium) (Thieu et al., 2009), and the Seine (France) ~~Garnier et al. (2020)~~ (Garnier et al., 2020) where discrepancies were noticed between the modeled and observed DO. Yang et al. (2010) used the WASP model to estimate DO in low-flow streams and noted that the uncertainty
45 of the model lies in the difficulty to characterize accurately the OM degradation and nitrification rates. The QUAL2E-OTIS water quality model shows similar discrepancies, which led Bailey and Ahmadi (2014) to conduct a sensitivity analysis to identify the governing parameters on DO. The ProSe model, which will be used in this study, also has mismatches at low flow ~~(Vilmin et al., 2018; ?; Garnier et al., 2020)~~ (Vilmin et al., 2018; Garnier et al., 2020; Wang et al., 2022). Moreover, Cox (2003) compares the existing water quality models on lowland rivers (flowing slowly with low DO content) and emphasizes
50 that the existing models lack one or more of the required criteria, in particular the inability to account for the expected un-certainties. ~~Thus, by observing the existing models in the market, it can be understood that large discrepancies exist between DO simulations and observations during low-flow periods and that most models are not able to simulate the evolution of DO accurately.~~ On the other hand, Wang et al. (2019b, 2022) assume that the uncertainties related to parameterization of (i) OM degradation kinetics and (ii) repartition of OM input from tributary rivers, WWTPs, and CSOs among dissolved and par-
55 ticulate pools ~~are assumed to~~ play a role in the discrepancies during non-bloom low-flow periods ~~(Wang, 2019a; ?; ?)~~, without explicitly quantifying their relative influences. However, in this paper, parameters representing OM kinetics and OM repartition are included with the ones of Wang et al. (2018) to quantify their impact on DO variation.

Therefore, a sensitivity analysis (SA) is necessary to study the role of the organic matter contributed by the model boundary conditions on the evolution of DO and river metabolism during low-flow periods via a parameterization of the organic matter

repartition into biodegradable fractions and its degradation by bacterial decomposition. Several applications of SA methods can be found for water quality modeling as in Nossent et al. (2011) and for DO and NO_3 modeling as in Bailey and Ahmadi (2014). Wang et al. (2018) summarized a list of SA applications in hydrological and water quality modeling and applied SA in contrasting hydrological and trophic contexts, where bacterial parameters were identified as the most influential in a 80-h non-bloom low-flow period. However, the sensitivity of water quality models has not been investigated for long-term (= 45 days to be consistent with the batch test for biodegradable fractions of OM (Servais et al., 1995)) low-flow periods against new parameters accounting for boundary condition uncertainties (OM repartition and degradation) as well as against bacterial and physical parameters.

This paper investigates not only the influence of bacterial properties on DO evolution at low flow, but also, for the first time, the role of the quantity and different fractions of OM sources on DO evolution. To further understand its functioning, the intra-parameter interactions are considered to evaluate the hiding effect of one parameter on the others. Moreover, a long period simulation (45 days) is conducted that leads to a better understanding of the mid-term effect of slowly biodegradable OM.

Based on the above discussion, the following three research questions are proposed to be answered in this study

1. What are the influential parameters controlling DO during a post-bloom summer low flow period where discrepancies are observed in different water quality models? Is a model that includes bacteria physiological parameters only sufficient to describe DO variation?

2. To what extent is the knowledge of the quantity of OM share, especially that of the biodegradable fraction of dissolved organic matter (BDOM) influential for water quality modeling?

3. What is the hierarchy among the influential parameters ?

In this article, the sensitivity of C-RIVE (Vilmin et al., 2012; Wang et al., 2018), the biogeochemical module of ProSe-PA (?)(Wang et al., 2019b), is investigated against the background of the aforementioned parameters based on 45- and 5-day DO simulations. After incorporation of new parameters to account for repartition and degradation of OM, SA is conducted on a synthetic river system mimicking the Seine. Finally, the parameters influencing the evolution of DO and those governing the degradation and repartition of OM are selected. On the basis of the results, proposals are made for a better integration of the influential parameters in data assimilation where the model is coupled with observation data to make an optimal estimate of the temporal evolution of the parameters as well as to produce better simulation results (Cho et al., 2020). Finally, some suggestions are made for water quality monitoring in urban areas in order to fulfill the modeling and monitoring requirements.

2   Material and methods

This section describes the forward model, the new parameterization and the bibliographically reviewed variation ranges, the SA strategy, and the simulations settings of the study. Since the main goal is to use a SA method to determine the controlling parameters that influence DO evolution during a summer low-flow period, we ~~considered~~ consider the C-RIVE model (section

2.1.1) as the forward model of the study and ~~identified~~ identify the parameters that need to be included in the study. Then, two new sets of parameters ~~were~~ are added to the study to account for the uncertainties related to the parameterization of OM degradation kinetics and its repartition into different constituent fractions (section 2.2). This ~~was followed by determining~~ is

95   followed by the determination of the variation ranges of the introduced parameters (2.2.3).

Three different SAs ~~were~~ are carried out on the basis of a specific SA strategy detailed in section 2.4 based on which 260K - 360K simulations ~~were~~ are run with the forward model on a case study resembling a non-bloom low flow (section 2.3). Depending on the output of these simulations that are DO time series, the Sobol SA method (section 2.5) ~~was~~ is implemented to determine the influential parameters. The Sobol indices ~~were~~ are calculated up to the second order so as to observe the

100   inter-parameter interactions in addition to their direct and total effects.

2.1 ProSe-PA

ProSe-PA results from the coupling of the ProSe hydro-biogeochemical model with a particle filter ~~(??)~~(Wang et al., 2019b, 2022) . It was developed to assimilate high-frequency observation data for a better estimation of the model parameters and an improvement in ProSe simulation results. The ProSe model ~~is~~ has been developed to simulate the hydro-biogeochemical evolu-

105   tion of the Seine from upstream of Paris until Poses (close to the estuary) and has been applied and validated numerous times ~~(Even et al., 1998, 2007; Flipo et al., 2004, 2007; Polus et al., 2011; Raimonet et al., 2015; Vilmin et al., 2015b, a; ?)~~(Even et al., 1998, 20 . In this model, the river is divided into longitudinal cells of specific length, where three ~~set~~ sets of equations corresponding to the three modules of ProSe-PA are solved (hydrodynamics, transport, and biogeochemical). First, the hydrodynamic equations are used to determine the discharge, velocity, and depth at each cell and at each time step, followed by the transport equations

110   for advection, dispersion, and diffusion, and finally the biogeochemical RIVE model equations for the concentrations of all the dissolved and particulate matter.

2.1.1 C-RIVE Biogeochemical model

C-RIVE is a C ANSI library that implements RIVE (Billen et al., 1994) concepts. It is the biogeochemical module of ProSe-PA ~~(?)~~(Wang et al., 2019b), which simulates the cycles of carbon, O_2, and other nutrients both in the water and sediment columns

115   of the river. The exchange of dissolved and particulate material between the two layers occurs through diffusion (Boudreau, 1997) and sedimentation-resuspension (due to shear flow and river navigation) (Martin, 2001; Even et al., 2004; Vilmin et al., 2015b), respectively. Numerous applications of RIVE exist for the ProSe and RIVERSTRAHLER softwares (See for instance Garnier et al. (1995, 2005); Even et al. (1998); Flipo et al. (2004, 2007); Thieu et al. (2009); Vilmin et al. (2016); Marescaux et al. (2020)).

120   RIVE simulates the macro-nutrients cycles (C,N,P,O_2) based on the physiology of micro-organisms living in water and/or sediments (heterotrophic bacteria, nitrifying bacteria, and phytoplankton), the kinetics of underlying physical-chemical processes, and carbon and nutrient inputs. Model parameters (a ~~hundreds~~ hundred to a few hundreds depending on the number of mico-organisms' species simulated) are mostly determined through experiments. Wang et al. (2018) summarized the list of physical, bacterial, and phytoplanktonic parameters related to the carbon cycle with their variation ranges. We examine the

equations for DO and OM evolution to understand the role of the different parameters and to select the appropriate parameters for inclusion in the SA.

2.1.2 Dissolved oxygen evolution equations

DO in the water column (Fig. 1) depends on physical, bacterial, and phytoplanktonic processes:

$$\frac{d[O_2]}{dt} = \frac{d[O_2]}{dt}_{physical} + \frac{d[O_2]}{dt}_{phytoplanktonic} + \frac{d[O_2]}{dt}_{bacterial} \tag{1}$$

The physical process depends on reaeration due to dams, wind, navigation, the oxygen holding capacity of water, and the diffusion of oxygen between the water-sediment interface as follows:

$$\frac{d[O_2]}{dt}_{physical} = \frac{K_{rea}}{h}([O_2]_{sat}(T) - [O_2]) - \frac{D_s}{h}([O_2]_{water} - [O_2]_{sed}) + \frac{d[O_2]}{dt}_{dams} \tag{2}$$

where,

h: water depth [m]

$[O_2]_{sat}(T)$: the saturated oxygen concentration at temperature T $[mgO_2/L]$

D_s: the coefficient of diffusion between water and sediment layer $[m/s]$

K_{rea} : the reoxygenation coefficient calculated from the empirical formula of Thibodeaux et al. (1994) as follows:

$$K_{rea} = \sqrt{\frac{D_m V_{wat}}{h}} + (K_{wind}V_{wind}^{2.23}(D_m * 10^4)^{2/3} + K_{navig}) \tag{3}$$

where,

K_{wind}: reoxygenation coefficient due to wind $[m/s]$

V_{wind}: wind speed $[m/s]$

V_{wat}: river flow velocity $[m/s]$

K_{navig}: reoxygenation coefficient due to navigation $[m/s]$ (Vilmin, 2014)

D_m: molecular diffusivity of DO $[m^2/s]$

The phytoplanktonic process depends on phytoplankton respiration ($R_{O_2,pp}$) and photosynthesis ($P_{O_2,pp}$) as follows:

$$\frac{d[O_2]}{dt}_{phytoplanktonic} = P_{O_2,pp} - R_{O_2,pp} \tag{4}$$

And the bacterial process that is the main source of oxygen consumption depends on the heterotrophic bacterial kinetics and the availability of substrate matter (S, considered to be the rapidly biodegradable dissolved organic matter, DOM_1, in this model) as follows:

$$\frac{d[O_2]}{dt}_{bacterial} = -\tau_{HB}(1 - Y_{HB}) \text{ uptake} \tag{5}$$

$$uptake = \frac{\mu_{max,HB}e^{\frac{-(T-T_{opt,HB})^2}{\sigma_{HB}^2}}\frac{[S]}{[S]+K_S}[HB]}{Y_{HB}} \tag{6}$$

where,

$[HB]$: the concentration of heterotrophic bacteria (hereafter, called bacteria) $[mgCL^{-1}]$

τ_{HB}: 1.0 $[molO_2/molC]$ for full oxidation of OM in the respiration process

Y_{HB}: the growth yield of heterotrophic bacteria [-]

$uptake$: the uptake of substrate (here S = DOM_1) for bacteria growth $[mgCL^{-1}h^{-1}]$

$T_{opt,HB}$: optimum temperature for the growth of bacteria $[^{\circ}C]$

$\mu_{max,HB}$: the maximum growth rate of bacteria at $T_{opt,HB}$ $[/h]$

σ_{HB}: standard deviation of bacteria temperature function $[^{\circ}C]$

K_s: Monod half-saturation constant for bacterial growth (uptake constant) $[mgCL^{-1}]$

2.1.3 Organic matter degradation equations

The OM in the Seine originates from (i) point releases from WWTP and CSO, (ii) diffuse sources due to soil leaching and surface degradation through tributary rivers, and (iii) mortality of bacteria and phytoplankton (Billen et al., 2001). Figure 1 illustrates the OM related process of C-RIVE in the water column where the total organic matter (TOC) is initially divided into dissolved (DOM) and particulate (POM) forms. DOM is composed of three different biodegradable fractions of (i) DOM_1 as the limiting substrate (rapidly biodegradable DOM in 5 days), (ii) DOM_2 (slowly biodegradable DOM in 45 days), and (iii) DOM_3 (refractory DOM). Similarly, POM is composed of (i) POM_1 (rapidly biodegradable POM), (ii) POM_2 (slowly biodegradable POM), and (iii) POM_3 (refractory POM). The benthic processes are not presented in Fig. 1.

The degradation of OM happens through the uptake of small monomeric organic substrates (S, here $S = DOM_1$) by heterotrophic bacteria on the basis of the HSB model (Billen et al., 1988; Servais, 1989; Billen, 1991) and presented by Eq. (7) and (9). These substrates are either the direct input (P_S) of DOM_1 from OM sources or produced from the exoenzymatic hydrolysis of the macromolecular fractions of both dissolved (DOM_2) and particulate (POM_1, POM_2) organic matter (Billen, 1991) or they originate from the phytoplankton excretion, which produces more easily utilizable OM (DOM_1) and microorganism lysis products that are macromolecular matter (Fig. 1) (Larsson and Hagstrom, 1979; Garnier and Benest, 1990; Billen, 1991).

$$\frac{d[S=DOM_1]}{dt} = hyd_{DOM_2} + hyd_{POM_{1,2}} - uptake_{HB} + P_S + P_E + P_L \tag{7}$$

where,

P_S, P_E, P_L: represent DOM_1 from the direct input of OM sources, phytoplankton excretion, and microorganism lysis, respectively $[mgCL^{-1}h^{-1}]$
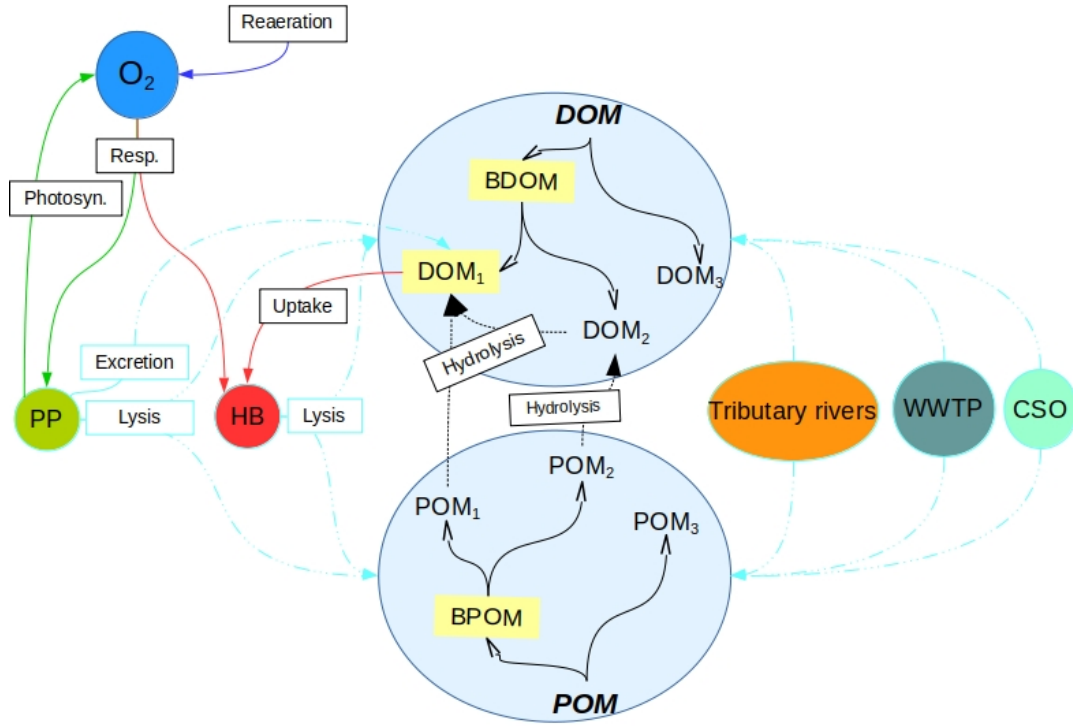
Figure 1. Schematic description of the OM-related process of C-RIVE in the water column. POM: particulate organic matter; DOM: dissolved organic matter; BDOM:biodegradable DOM, BPOM:biodegradable POM (subscripts 1, 2, and 3 refer to rapidly degradable, slowly degradable, and non-biodegradable fractions of OM, respectively); Cyan ~~solid~~ dashed-dotted lines: OM input from sources and repartition between POM and DOM; Solid black lines: repartition of DOM and POM into biodegradability pools ; Dotted black lines: Hydrolysis; Remaining solid lines: Biogeochemical processes. Resp.:Respiration; Photo,:Photosynthesis; PP: primary producer; HB: heterotrophic bacteria; WWTP: wastewater treatment plant; CSO: combined sewage overflow

180 hyd_{DOM_2}: hydrolysis of DOM_2 into DOM_1 based on the exoenzymatic hydrolysis equation of Michaelis-Menten $[mgCL^{-1}h^{-1}]$

$hyd_{POM_{1,2}}$: hydrolysis of POM_1 and POM_2 into DOM_1 and DOM_2, respectively, by first-order kinetics $[mgCL^{-1}h^{-1}]$

$$hyd_{DOM_2} = k_{hyd,max} \frac{[DOM_2]}{[DOM_2] + K_{DOM2}}[HB] \tag{8}$$

$$uptake_{HB} = \mu_{max,HB} \frac{[DOM_1]}{[DOM_1] + K_s}[HB] \tag{9}$$

185 where,

$uptake_{HB}$: uptake or consumption of DOM_1 by heterotrophic bacteria $[mgCL^{-1}h^{-1}]$

$k_{hyd,max}$: coefficient for hydrolysis of DOM_2 into DOM_1 [$/h$]

K_{DOM2}: constant of semi-saturation for the hydrolysis of DOM_2 [$mgCL^{-1}$]

2.2 Parameterization of organic matter share (repartition) and degradation

190 In order to account for the uncertainties related to the parameterization of OM degradation kinetics and its repartition into different constituent fractions, the following two sets of parameters are introduced:

2.2.1 OM degradation parameters

C-RIVE The parameters related to OM degradation are K_s (represents uptake of DOM_1 by bacteria), K_{DOM2} and $k_{hyd,max}$ (represent hydrolysis of DOM_2 to DOM_1), which have been defined in section 2.1.3 and that already exist in C-RIVE.
195 Hydrolysis parameters of POM are not considered in this study because the rate of hydrolysis of $POM_{1,2}$ is slower than that of DOM_2 by an order of magnitude of 100 to 1000 (Billen et al., 1994).

2.2.2 OM repartition or share parameters

The OM repartition parameters are a novelty added in C-RIVE does. Indeed, former version of C-RIVE did not include any parameter to define the repartition of OM into DOM and POM and then further into their corresponding fractions $DOM_{1,2,3}$
200 and $POM_{1,2,3}$; therefore, the. The following five parameters are introduced to represent the repartition of OM:

$$t = \frac{DOM}{TOC}$$

$$b_1 = \frac{BDOM}{DOM}$$

205 $$s_1 = \frac{DOM_1}{BDOM}$$

$$b_2 = \frac{BPOM}{POM}$$

$$s_2 = \frac{POM_1}{BPOM}$$

210 where,

TOC: total organic matter or carbon (= $DOM + POM$) [$mgCL^{-1}$]

$BDOM$: biodegradable DOM (= $DOM_1 + DOM_2$) [$mgCL^{-1}$]

$BPOM$: biodegradable POM (= $POM_1 + POM_2$) [$mgCL^{-1}$]

t: ratio between dissolved and total organic matter [-]

215 b_1: ratio between biodegradable DOM and DOM [-]

s_1: ratio between rapidly biodegradable DOM and biodegradable DOM [-]

b_2: ratio between biodegradable POM and POM [-]

s_2: ratio between rapidly biodegradable POM and biodegradable POM [-]

However, these parameters are not inserted into C-RIVE; instead, the six OM fractions (To further clarify, $DOM_{1,2,3}$ and $POM_{1,2,3}$, which are were state variables in the former version of C-RIVEinputs) are calculated in terms of the aforementioned five independent parametersbased on the classification presented in Fig. 2 and thereby they are reflected in the six model inputs. . They used to be forced information provided by user. They are now defined by the proposed repartition model which has the above-mentioned five parameters. The only required forced information is TOC that comes from experimental data. Therefore, the five OM repartition parameters give us the possibility to conduct a sensitivity analysis to quantify their influence on the output of C-RIVE namely DO concentration. Thanks to this model, we can have varying $DOM_{1,2,3}$ and $POM_{1,2,3}$ depending on the values of the five parameters in a specific river.
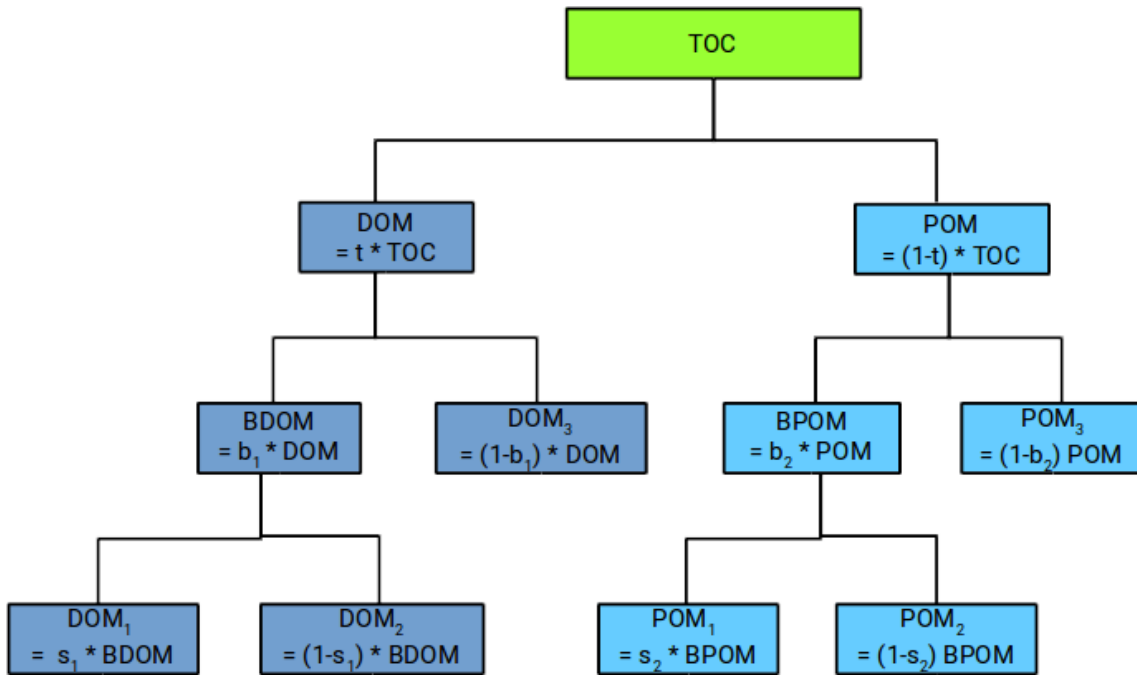
Figure 2. OM repartition into the six fractions of dissolved and particulate matter in terms of the five repartition parameters, namely, t, b_1, s_1, b_2, s_2

2.2.3 Parameters for SA and bibliographical review of their variation ranges

Following the parameterization, the 17 parameters as listed in Table 1 are evaluated in the SA. This includes two physical parameters that account for O_2 re-aeration; seven bacterial parameters that account for bacteria growth, mortality, and respiration; three OM degradation parameters that demonstrate OM kinetics; and five OM share parameters that represent the repartition of TOC into smaller dissolved and particulate fractions.

Before proceeding to SA, it is necessary to specify the range of variation of these parameters according to the existing literature. As indicated in Table 1, the range of variation of the repartition and degradation parameters is selected based on a detailed bibliographical review, which is discussed and tabulated thoroughly in Hasanyar et al. (2020, 2021) , while the variation of physical and bacterial parameters is retrieved from Wang et al. (2018). Table 1 also includes the range of variation of TOC, which represents the total organic matter input in the model due to the boundary conditions and varies from 1 to 10 mgCL^{-1} under low flow (also retrieved from Hasanyar et al. (2020)). However, phytoplanktonic parameters are not included in this study because they exert an influence only during algal bloom periods (Wang et al., 2018), whereas this study is conducted under non-bloom situations where heterotrophic activity is dominant.

2.3 Case study

The synthetic case developed by Wang et al. (2018) ~~was~~ is adapted for the application of SA methods on C-RIVE parameters during a low-flow period (Fig. 3). It is a stretch with a width of 100 m and a length of 1000 m representing the Seine. The low-flow period is identified with a discharge of 80 m^3/s based on the data at Bougival station during the summer season. The simulation period is set at 45 days in order to be coherent with the experimental protocol of the BDOM measurement (Servais et al., 1995) where it is considered as the threshold between the biodegradable and refractory fractions of TOC in a batch experiment. On the other hand, a 45-day simulation period is also necessary for studying the long-term effect of TOC degradation.
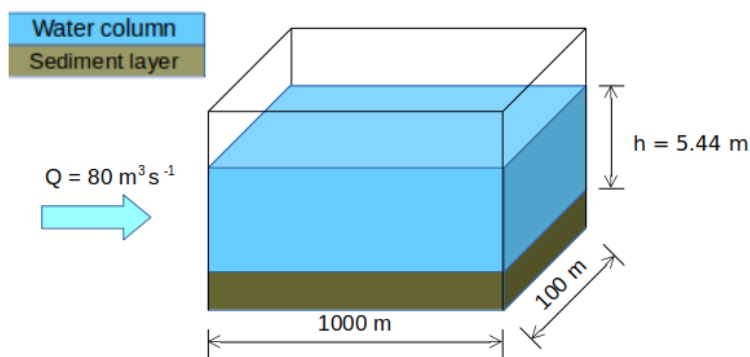
Figure 3. Synthetic scheme representing a reach of the Seine (modified from Wang et al. (2018))

2.3.1 Initial conditions

A non-bloom low-flow situation (large heterotrophic bacteria biomass presence in the river) is considered to represent the low-flow period. Table 2 lists the initial concentrations for both water and sediment compartments that are set based on the mean

Table 1. list of parameters and their corresponding ranges of variation

Parameter	Description	Min. Val.	Max. Val.	Unit	References
TOC	Total organic carbon	1	10	[mgCL^{-1}]	
OM share parameters					
t	ratio between dissolved and total organic matter (DOM/TOC)	0.4	0.9	[-]	
b_1	ratio between biodegradable DOM and DOM (BDOM/DOM)	0.1	0.5	[-]	
b_2	ratio between biodegradable POM and POM (BPOM/POM)	0.1	0.5	[-]	
s_1	ratio between rapidly biodegradable DOM and BDOM (DOM_1/BDOM)	0.4	0.95	[-]	Hasanyar et al. (2020, 2021)
s_2	ratio between high biodegradable POM and BPOM (POM_1/BPOM)	0.4	0.95	[-]	
OM degradation parameters					
K_s	constant of semi saturation for bacterial substrate uptake	0.02	0.15	[mgCL^{-1}]	
K_{DOM2}	constant of semi saturation for the hydrolysis of DOM_2	0.2	1.5	[mgCL^{-1}]	
$k_{hyd,max}$	coefficient of the hydrolysis of DOM_2 to DOM_1	0.25	0.75	[/h]	
Bacterial parameters					
$T_{opt,hb}$	optimum temperature for bacterial growth	15	30	[°C]	
σ_{hb}	standard deviation of temperature function for bacterial growth	12.75	21.25	[°C]	
$V_{sed,hb}$	settling velocity of bacteria	0	0.1	[m/h]	
$K_{O_2,hb}$	Half-saturation constant for dissolved oxygen	0.375	0.625	[mgO_2/L]	Wang et al. (2018)
$\mu_{max,hb}$	maximum growth rate of bacteria	0.01	0.07 [*]	[/h]	
Y_{hb}	bacterial growth yield	0.03	0.5	[-]	
mort$_{hb}$	bacterial mortality rate	0.01	0.08	[/h]	
Physical parameters					
K_{navig}	re-aeration coefficient due to navigation	0	0.05	[m/h]	
K_{wind}	re-aeration coefficient due to wind	0.885	1.475	[m/h]	

[*] The upper limit identified by Wang et al. (2018) is decreased from 0.13/h to 0.07/h in order to avoid complete DO depletion in simulations longer than 5 days

concentrations of the simulations at Bougival station during the 2007-2012 period (Vilmin et al., 2016) except for temperature (depending on summer season), DO (depending on oxygen solubility), POM and DOM fractions (depending on the TOC concentration and share parameters), and phytoplankton and bacterial biomass (depending on a post-bloom condition). Indeed, the sum of $DOM_{1,2,3}$ and $POM_{1,2,3}$ is equal to the desired TOC, but they are distributed among the six fractions based on the

five OM share parameters (t, b_1, s_1, b_2, s_2). The hyporheic exchanges (between groundwater and river) are ignored in this work because the contribution of groundwater to downstream rivers is well known to be negligible with respect to the discharge of large rivers (Strahler order > 6) (Pryet et al., 2015).

Table 2. Initial concentrations of the simulations

No	Species	$C_{ini,water}$	$C_{ini,sediment}$	Unit
1	NH_4	0.12	0.33	[mgN/L]
2	NO_2	0.04	0.04	[mgN/L]
3	NO_3	7	4.54	[mgN/L]
4	TSS	16.82	95010	[mg/L]
5	PO_4	0.1	0.27	[mgP/L]
6	O_2	8.62	6.65	[mgO_2/L]
7	HB	0.023	0.016	[mgCL^{-1}]
8	PP	0.010	0.003	[mgCL^{-1}]
9	DOM_1	f(TOC, share parameters)	0.12	[mgCL^{-1}]
10	DOM_2	f(TOC, share parameters)	1.28	[mgCL^{-1}]
11	DOM_3	f(TOC, share parameters)	1.94	[mgCL^{-1}]
12	POM_1	f(TOC, share parameters)	44	[mgCL^{-1}]
13	POM_2	f(TOC, share parameters)	696	[mgCL^{-1}]
14	POM_3	f(TOC, share parameters)	2555	[mgCL^{-1}]
15	T_{mean}	22.4 ± 3.0		°C

2.4 Sensitivity analysis strategy

260 The objective is to identify different scenarios under which the SA has to be conducted in order to detect the influential parameters under a summer low-flow condition. Therefore, a coarse pre-analysis consisting in forward simulations of the C-RIVE model is first conduced with extreme values of a small number of representative parameters. Then, the necessary scenarios are developed to assess the assumptions and conclusions put in place in the pre-analysis.

2.4.1 Pre-analysis of the model under parameter extreme limits

265 First, we need to select certain parameters for the pre-analysis. We consider $\mu_{max,hb}$, $mort_{hb}$ and Y_{hb} as they were found to be influential in the study of Wang et al. (2018) under non-bloom situations. However, to decrease the number of parameters, $mort_{hb}$ and $\mu_{max,hb}$ are represented together as a single parameter called "net growth (NG)."

$$Net\ Growth\ (NG) = \mu_{max,hb} - mort_{hb}$$

Fixing $mort_{hb}$ = 0.02/h at its reference value and $\mu_{max,hb}$ ranging between 0.022 and 0.07/h, net growth was found to range from 0.002 to 0.05/h while the range for Y_{hb} is taken from Table 1. On the other hand, as the OM share parameters are not C-RIVE inputs, we consider BDOM to represent them in the model. Its range is calculated in Eq. (10)-(11) in order to be statistically independent based on the TOC repartition diagram (Fig. 2) as follows:

$$BDOM_{min} = TOC_{ref} * t_{ref} * b_{1,min} \tag{10}$$

$$BDOM_{max} = TOC_{ref} * t_{ref} * b_{1,max} \tag{11}$$

Here, TOC_{ref} is a reference TOC value and fixed at 5 mgCL^{-1} (considered as the baseline concentration of TOC in the Seine (Vilmin et al., 2016)), the reference t (t_{ref}) = 0.7 is the average value of t variation range and b_1 is taken from Table 1.

Table 3. Combinations of the three parameter values for eight single simulations

Sim. No.	BDOM	Net growth	Y_{hb}
1	0.35	0.05	0.03
2	0.35	0.002	0.03
3	1.75	0.05	0.03
4	1.75	0.002	0.03
5	1.75	0.05	0.5
6	1.75	0.002	0.5
7	0.35	0.05	0.5
8	0.35	0.002	0.5

Therefore, eight simulations pertaining to eight different combinations of the minimum and maximum values of these three parameters are conducted (Table 3) and accordingly for each combination, the evolution of DO, DOM_1, DOM_2, and BDOM is plotted (Fig. 4). In order to discriminate easily among the eight plots of single simulations using their titles, any parameter name written in capital or small letters means that its maximum or minimum value is used, respectively in that particular simulation. For example, plot 3 (BDOM NG y), that corresponds to simulation 3 in Table 3, is a simulation where the maximum values of BDOM and net growth and the minimum value of Y_{hb} are used.

As can be observed in Fig. 4, simulations 3 (BDOM NG y), 4 (BDOM ng y), 5 (BDOM NG Y) and 6 (BDOM ng Y) have high BDOM where all except simulation 6 demonstrate considerable DO depletion. This shows the importance of BDOM in the depletion of DO. The reason for lack of depletion in 6 could be attributed to the combination of low net growth and a high yield due to which BDOM is not consumed. However, comparing simulations 6 (BDOM ng Y) and 7 (bdom NG Y), we observe that even a lower BDOM coupled with high net growth (simulation 7) has more effect on DO than a high BDOM coupled with
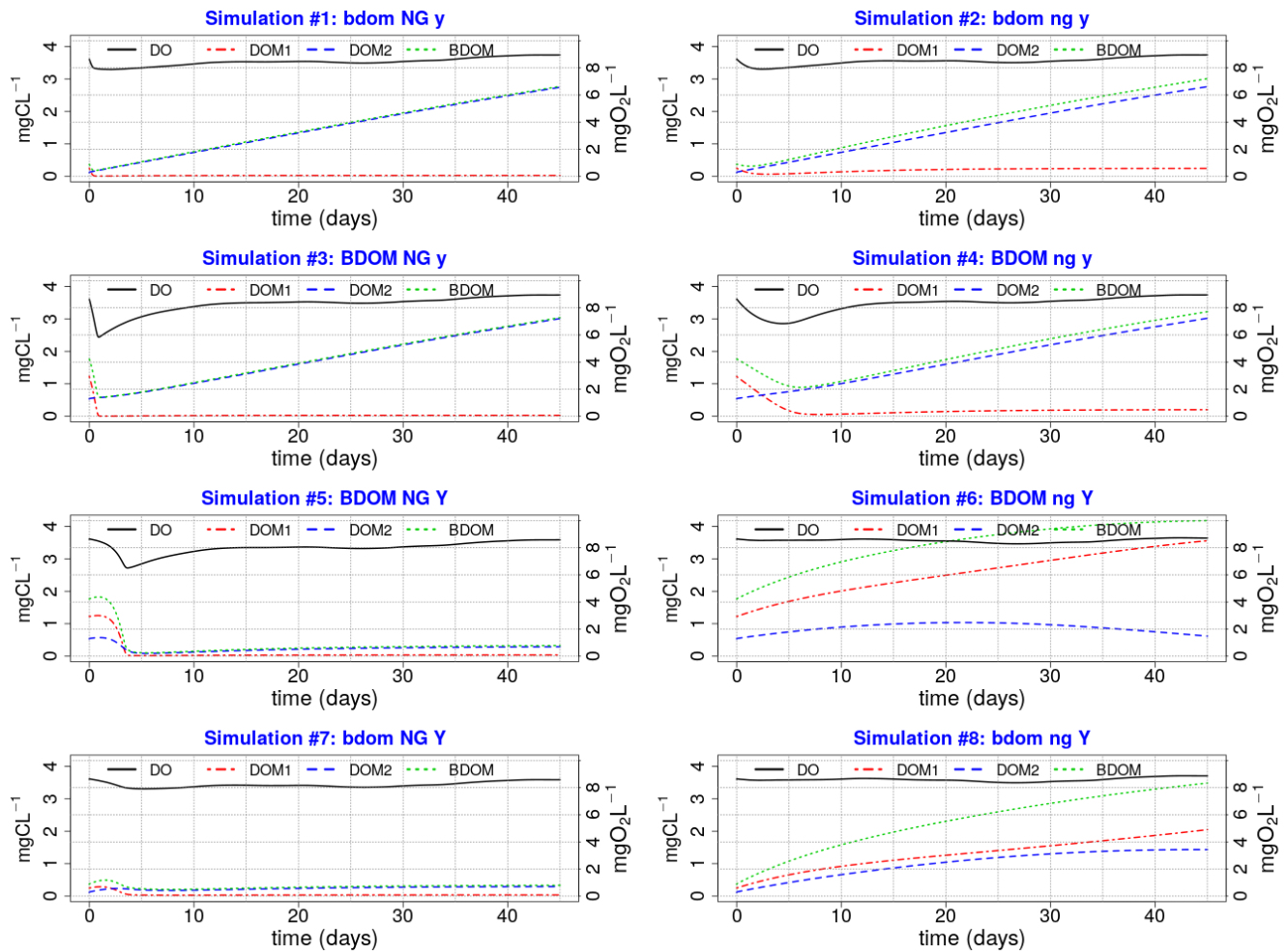
Figure 4. Eight plots of single simulations; (XX yy ZZ) Any parameter name written in capital or small letters means that its maximum or minimum value is used, respectively, in that specific single simulation. For example, plot #3 (BDOM NG y) that corresponds to simulation #3 in Table 3 is a simulation where the maximum values of BDOM and net growth and the minimum value of Y_{hb} are used

low net growth (simulation 6). This shows the interaction effect of BDOM with net growth parameters and the fact that despite BDOM being the primary requirement for depletion of DO, the net growth needs to be high in order to demonstrate the influence of BDOM and provide the means for its consumption which would result in DO depletion. Moreover, the difference between two horizontally adjacent simulations is in the net growth, which is maximum for the left-hand simulations and minimum for those on the right and as a result, all simulations on the left side with a high net growth demonstrate more depletion or consumption of DO than those on the right side. This shows the influence of net growth parameters ($\mu_{max,hb}$, $mort_{hb}$) on the model at low flow.

2.4.2 Sensitivity analysis scenario development

Having performed the pre-analysis, we understood the importance of BDOM and that of the net growth parameters such that BDOM is needed primarily in order to be consumed so that DO could be depleted. And then we discern that BDOM consumption is high when net growth is at its highest value. Therefore, the following three different sensitivity analysis scenarios whose parameters are detailed in Table 4 need to be conducted.

The first SA is conducted by assuming the general influence of net growth parameters in the pre-analysis, and in order to have a broader view of the model sensitivity with respect to all the model parameters. Based on the pre-analysis, we observed that the main effect due to BDOM is linked to high net growth rates, therefore, we can assume that the effect of parameters other than net growth parameters is demonstrated when they are coupled with a high net growth condition. In addition, since a significant interaction (the difference between the first and total sensitivity indices) is observed between net growth parameters in Wang et al. (2018), they are assumed to be hiding the influence of other parameters. Therefore, in order to confirm these two assumptions and to observe the influence of other parameters, we implement a second SA where net growth parameters are fixed at its highest value. This SA removes the possibility of interaction among net growth and other model parameters. It results in a better evaluation of the model sensitivity with respect to the parameters whose influences might be hidden by the dominant and interacting parameters.

The third SA is performed to verify the second SA assumption that parameters other than net growth exert their influence only under a high net growth condition, and thus the same parameters could be deemed non-influential under a low net growth situation. therefore the net growth parameter is fixed at its lowest value. The settings for the three SAs are as follows:

1. **First SA (All parameters included)**: There are 17 defined parameters (Table 1 & Table 4) in the model, the simulation period is **45 days**. It is repeated 10 times for every TOC concentration of 1-10 mgCL^{-1}.

2. **Second SA (Fixed high net growth)**: The net growth parameters are fixed as follows:

$$Fixed\ high\ net\ growth = High\ bact.\ growth\ rate\ (\mu_{max,hb} = 0.07/h) - Bact.\ mortality\ rate\ (mort_{hb} = 0.02/h)$$

Furthermore, to decrease the computational cost of the model, the three OM share parameters (t, b_1 and b_2) from the first SA are narrowed to BDOM and BPOM whose variation ranges are calculated based on the following Eq. (12)-(15) as follows:

$$BDOM_{min} = TOC * t_{ref} * b_{1,min} \tag{12}$$

$$BDOM_{max} = TOC * t_{ref} * b_{1,max} \tag{13}$$

$$BPOM_{min} = TOC * (1 - t_{ref}) * b_{2,min} \tag{14}$$

$$BPOM_{max} = TOC * (1 - t_{ref}) * b_{2,max} \tag{15}$$

Here TOC varies between 1-10, therefore, similar to the first SA, this SA is repeated 10 times corresponding to each TOC case. The pre-analysis also demonstrated that BDOM or precisely the substrate (DOM_1) is consumed in less than 5 days under the high net growth condition (simulations 1, 3, 5 & 7) , therefore imposing a **5 days** simulation period. Twelve parameters are evaluated under this scenario (Table 4).

3. **Third SA (Fixed low net growth)**: This SA is conducted in a similar way to the second SA except that this time $\mu_{max,hb}$ is fixed at a lower value of 0.022/h in order to simulate a very low net growth rate condition as follows:

$$Fixed\ low\ net\ growth = Low\ bact.\ growth\ rate\ (\mu_{max,hb} = 0.022/h) - Bact.\ mortality\ rate\ (mort_{hb} = 0.02/h)$$

Similar to the second SA, this SA is also repeated 10 times and implemented under a **5-day** simulation period. Twelve parameters are also evaluated under this scenario (Table 4).

2.5 Sensitivity analysis methodology

Each of the three aforementioned SAs is implemented based on an innovative SA methodology initially proposed in Wang et al. (2018) and adopted in this study, where the influence of input parameters (X) is evaluated on the C-RIVE model according to the variations of a large set of DO simulations (model output, Y). The followings steps are pursued in this approach:

1. **Input parameter identification**: Initially, a set of D input parameters (Table 4) are identified with their corresponding ranges of variation (Table 1).

2. **Parameter sampling and model input creation**: Saltelli's extension of the Sobol sequence (Saltelli, 2002) implemented in PYTHON SALIB package (Herman and Usher, 2017) is employed to create different combinations of the input parameters, which are designed to produce optimized simulations and efficient analysis results. Considering a sample size of 10,000 (N) (needed for stable results based on Nossent et al. (2011)), a matrix with a size of N(2D+2) $\times$ D is created for each SA scenario where every row represents one set of input parameters for the model.

3. **Model simulation**: In this step, the model inputs are launched into C-RIVE for the simulation period considered with a 1-min time step. As an output, a DO time series matrix with a size of $N(2D+2) \times M$, where M is the number of output time steps based on a 15 min output time step [1], is created corresponding to each input matrix created in the previous step. Figure 5 demonstrates the ensemble of 260,000 [=N(2D+2)] DO simulations of TOC = 5 mgCL^{-1} in the second SA scenario (TOC = 5 mgCL^{-1} is used in this study to represent the TOC range of 1-10 mgCL^{-1} and to show the results in case they are similar across all TOC concentrations).

---

[1]M = 45- or 5-day simulation period $\times$ 24 hrs $\times$ 3600 min / (1-min simulation time step $\times$ 15-min output time step) = 4230 or 480 depending on the simulation period, respectively

Table 4. The parameters considered in each of the four sensitivity analyses

	1st SA	2nd SA	3rd SA
OM share parameters	t	BDOM	BDOM
	b_1	BPOM	BPOM
	s_1		
	b_2		
	s_2		
OM degradation parameters	K_s	K_s	K_s
	K_{DOM2}	K_{DOM2}	K_{DOM2}
	$k_{hyd,max}$	$k_{hyd,max}$	$k_{hyd,max}$
Bacterial parameters	$T_{opt,hb}$	$T_{opt,hb}$	$T_{opt,hb}$
	σ_{hb}	σ_{hb}	σ_{hb}
	$V_{sed,hb}$	$V_{sed,hb}$	$V_{sed,hb}$
	$K_{O_2,hb}$	$K_{O_2,hb}$	$K_{O_2,hb}$
	Y_{hb}	Y_{hb}	Y_{hb}
	$\mu_{max,hb}$		
	$mort_{hb}$		
Physical parameters	K_{navig}	K_{navig}	K_{navig}
	K_{wind}	K_{wind}	K_{wind}
total number of parameters	17	12	12

4. **Dimensionality reduction**: The empirical orthogonal function (EOF) method is an adaptation of principal component
analysis (PCA) to study a phenomenon that changes with a continuous variable, such as time, and is applied to transform
the output data from one coordinate system into another by introducing new uncorrelated (orthogonal) variables (princi-
pal components) (Jolliffe and Cadima, 2016). EOF is adopted to transform the model output, which is a DO times series
matrix composed of M columns into a smaller matrix where each simulation can be represented by a linear combination
of EOFs. The coefficients of this linear combination are indeed orthogonal projections that maximize the variance while
transforming the data from a higher-dimensional space into a lower one. The way EOF decreases dimensionality is such
that it ranks the components based on the maximized variance. In other words, most of the information is kept in the first
few components, thereby making it possible to reduce the number of dimensions without losing a considerable amount
of information (Wold et al., 1987). In this study, the first k EOF elements that constitute at least 99% of the total model
variance are considered to represent each single simulation of the DO time series as shown for TOC = 5 mgCL^{-1} in the
second SA (Fig. 56a), where four (k) significant EOFs are found such that the first EOF (EOF_1) represents almost 55%
of the total variance. Figure 6b illustrates the evolution of the eigenvalues of the four (k) EOFs with time, which are
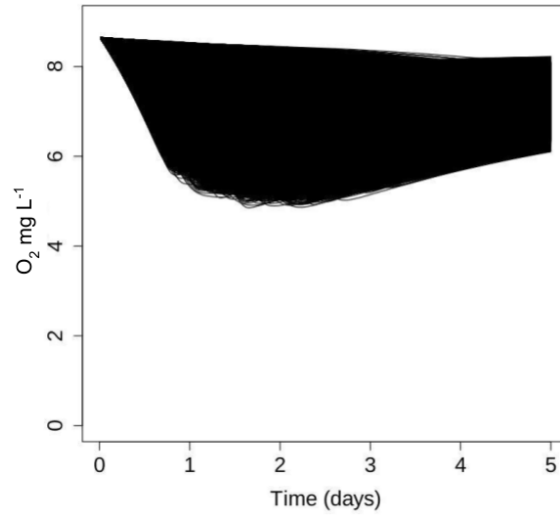
Figure 5. Ensemble of the 260,000 DO simulations for TOC = 5 mgCL^{-1} in the second SA scenario

consequently used to represent each simulation in terms of the k new coordinates . Thereby, an $N(2D+2) \times M$ matrix is converted into a new matrix of $N(2D+2) \times k$, which will be subjected to the Sobol SA method. The R $prcomp$ function is used to conduct the EOF analysis.
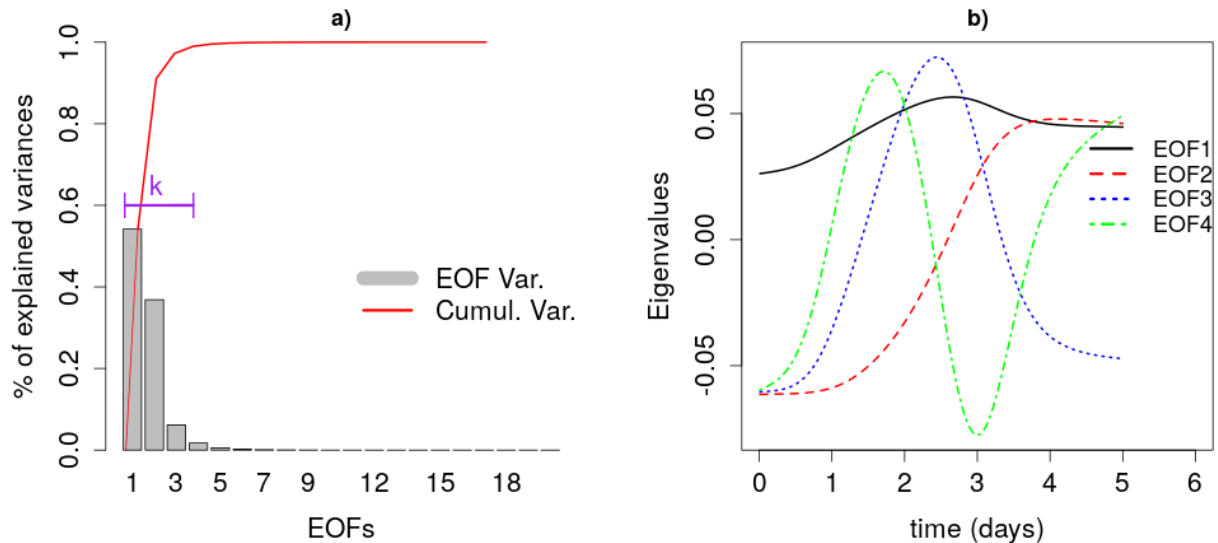
Figure 6. a) Cumulative sum of EOF variances and b) time evolution of four (k) significant EOFs for TOC = 5 mgCL^{-1} in second SA

5. **Sobol sensitivity analysis**: The Sobol SA method (Sobol, 1993; Saltelli et al., 2010) is applied in this study to evaluate the sensitivity of the model output against the input parameters. It is a variance-based method that classifies the parameters based on their contribution to and/or influence on the total variance of the model output (Brookes et al., 2015). It is a convenient method to be used for SA of complex models that involve interactions between parameters. In this method, the model output (Y) is expressed as a function of D parameters:

$$Y = f(X) = f(X_1, ..., X_D),$$ (16)

such that the model output could be decomposed by elementary functions:

$$f(X) = f_0 + \sum_{i=1}^{D} f_i(X_i) + \sum_{i=1}^{D-1} \sum_{j=i+1}^{D} f_{ij}(X_i, X_j) + ... + f_{1,...,D}(X_1, ..., X_D)$$ (17)

Here f_0 is the expectation of the model output and each one of the elementary functions have a zero mean and can be computed by integration:

$$f_0 = \int_{[0,1]^D} f(X)d_X$$ (18)

$$f_i(X_i) = -f_0 + \int_{[0,1]^{D-1}} f(X)d_{X \sim i}$$ (19)

$$f_{ij}(X_i, X_j) = -f_0 - f_i(X_i) - f_j(X_j) + \int_{[0,1]^{D-2}} f(X)d_{X \sim (ij)}$$ (20)

On the other hand, the total unconditional model variance could be defined as:

$$V(Y) = \int_{[0,1]^D} f^2(X)d_X - f_0^2$$ (21)

Thereby, the total unconditional variance of the model can be expressed as:

$$V(Y) = \sum_{i=1}^{D} V_i(X_i) + \sum_{i=1}^{D-1} \sum_{j=i+1}^{D} V_{ij}(X_i, X_j) + ... + V_{1,...,D}(X_1, ..., X_D)$$ (22)

where, V_i is the partial variance of the i_{th} parameter and V_{ij} is the interaction effect of the i_{th} and j_{th} parameters. The partial variance is calculated as:

$$V_{i_1,...,i_s} = \int_0^1 ... \int_0^1 f_{i_1,...,i_s}^2(X_{i_1}, ..., X_{i_s})dX_{i_1}, ...dX_{i_s}$$ (23)

where $s = 1, ..., D$ and f_i is an elementary function. Therefore, the first-order Sobol SA indices can be computed as follows:

$$S_i = \frac{V_i}{V} \qquad (24)$$

S_i is also called as the "main effect" because it represents the contribution of a single input parameter i on the total variance. The total sensitivity index (S_{Ti}), also called "global effect," is another index that represents the sum of the first-order index (S_i) and the effect of the interaction between the parameters and is calculated as follows:

$$S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + ... \qquad (25)$$

here, $S_{ij} = \frac{V_{ij}}{V}$ is called the "second-order index" and measures the interaction between a pair of parameters X_i and X_j. Therefore, the sum of second-order interactions of any parameter X_A with other parameters (X_B, ..., X_D) is considered to represent the second-order index of each parameter (S_2) as follows:

$$S_{2,A} = \sum_j S_{Aj} \qquad (26)$$

Since the output of previous step is a matrix of k vectors corresponding to the k EOFs, the Sobol indices of parameters are initially calculated k times for each EOF and then added while being weighted by the variance of the corresponding EOF.

The total computation time from step 1 to 5 is 12 h for each TOC case of the first SA, after using a parallel code in PYTHON and dividing the simulations in several groups to decrease the computational cost using 20 processors (Intel(R) Xeon(R) E5-2640 and frequency of 2.40GHz). The computation time is 3 h for each TOC case of the second and third SA, which clearly demonstrates the gain in time compared with the first SA.

3 Results

This section presents the results of the three Sobol SAs during a summer low-flow period. The influential parameters of each analysis are discussed in the following paragraphs.

3.1 First SA: All parameters

Figure 7a presents the results of the Sobol SA method on TOC = 5 mgCL^{-1}. It is expressed by a bar plot of the total sensitivity (S_T), first-order sensitivity (S_1), and second-order sensitivity (S_2) indices of the parameters. The higher-order sensitivity indices are also calculated in terms of the difference between the total and the first- and second-order indices (S_T-S_1-S_2). The parameters are ranked based on their S_T and the most influential parameters are shown by the shaded area, which includes

parameters constituting 95% of the total variance of the model output. However, this SA is conducted on different TOC values of 1-10 mgCL^{-1}, and therefore in order to summarize all of the results together, the evolution of the normalized total sensitivity indices (S_T^*) of the six most influential parameters with TOC is presented in Fig. 7b.
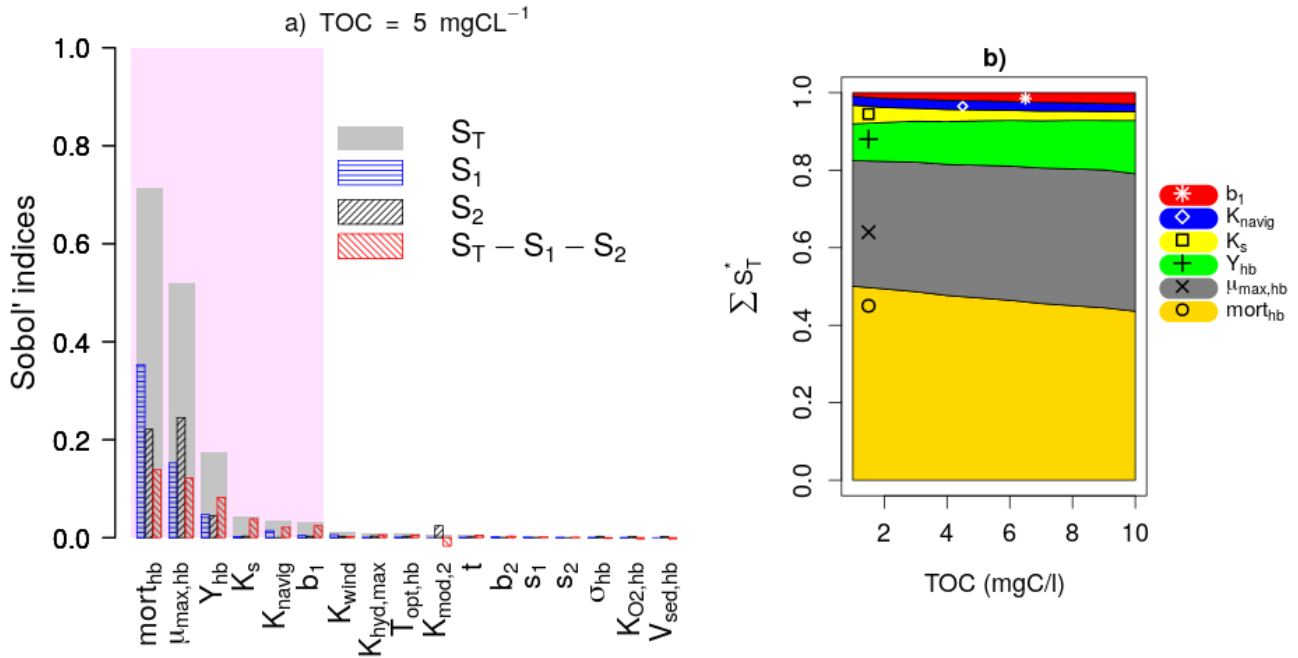
Figure 7. Sobol SA results of **first SA: All parameters** (a) Sobol SA results for TOC = 5 mgCL^{-1}; (b) Evolution of the normalized total sensitivity indices of the influential parameters with TOC

According to Fig. 7b, the three bacterial parameters of bacterial mortality rate ($mort_{hb}$), maximum bacterial growth rate ($\mu_{max,hb}$), and bacterial yield (Y_{hb}) exert the most influence on DO evolution, and whatever the TOC concentration, they represent the majority of the model sensitivity. By increasing TOC, we observe a gradual decrease in the influence of $mort_{hb}$, but an increase in the influence of Y_{hb}. This result obtained over the 45-day simulation period (Fig. 7) confirms the assumption made in the pre-analysis step (sec 2.4.1) regarding the overall dominance of bacterial parameters in long-term low-flow periods~~and is in accordance with the findings of the study by Wang et al. (2018), which was conducted over a 4-day simulation period~~.

Apart from the constant of navigation (K_{navig}), which is a physical parameter, the other two influential parameters (K_s, b_1) are OM-related parameters that were introduced in this study. K_s seems to be more important in lower TOC concentrations compared to b_1 whose influence increases in higher TOC concentrations.

On the other hand, these results confirm the assumptions made in the pre-analysis step that the dominant parameters tend to hide the influence of other parameters. Observing the second-order (S_2) and higher-order sensitivity (S_T-S_1-S_2) indices of $mort_{hb}$, $\mu_{max,hb}$ and Y_{hb} in Fig. 7a, very strong interactions can be highlighted between these parameters, i.e., a significant
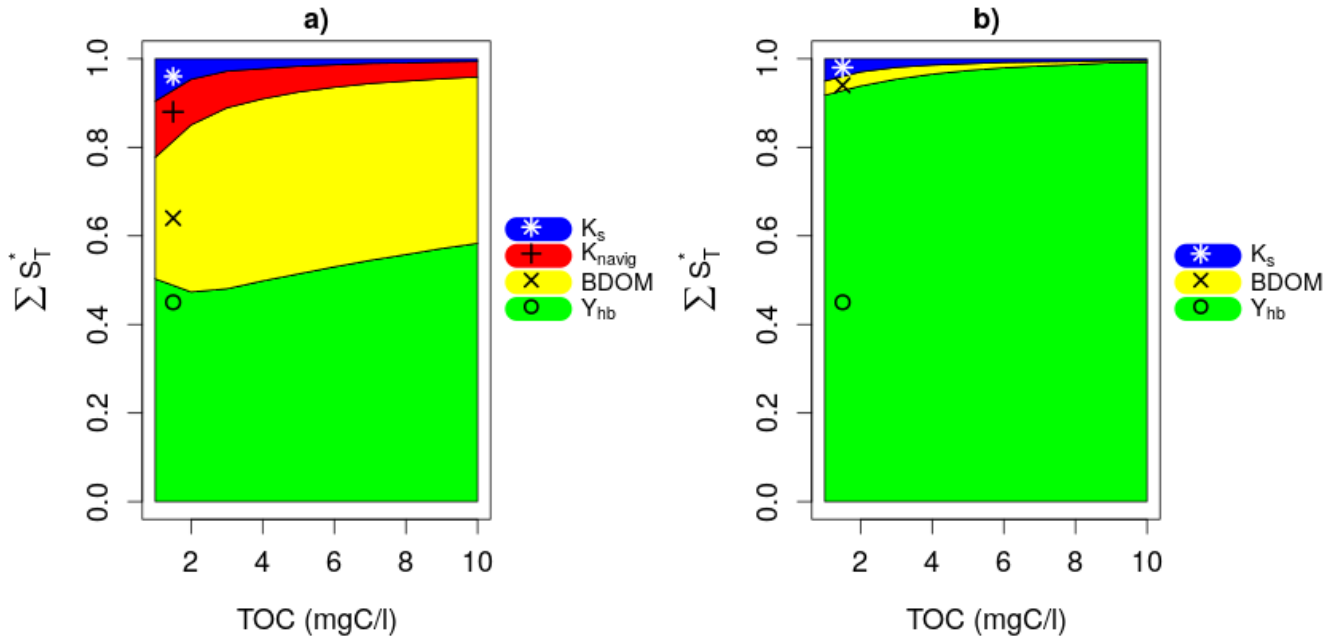
Figure 8. Results of (a) Second SA: Fixed high net growth; (b) Third SA: Fixed low net growth

portion of their total sensitivity indices is due to their internal interactions. On the other hand, b_1 and K_s also exert an influ-
435 ence due to their higher-order interaction with these three parameters. Thus, in order to see what happens behind net growth
parameters, and in order to observe the dominant parameters under two extreme conditions of high and low net growth rates,
the second and third SAs are conducted by fixing the $mort_{hb}$, $\mu_{max,hb}$ parameters under the notion of net growth rate.

Moreover, the three OM parameters (t, s_1 and s_2) are found to be non-influential, which means they can be excluded from SA
by fixation in the second and third SA while calculating the variation ranges of BDOM and BPOM (Eq. (10)-(13)). Finally, the
440 inclusion of b_1 among influential parameters out of the five OM parameters confirms the selection of the BDOM concentration
instead of other OM components in the pre-analysis step.

3.2 Second SA: Fixed high net growth

Based on Fig. 8a which depicts the results of the second SA, it is shown that Y_{hb} and BDOM are the most influential parameters
under high net growth rate conditions. This is due to the fact that the bacterial community manages to consume most of BDOM
445 under a high net growth condition and then at some point, BDOM becomes a limiting factor for their growth. This result
confirms the assumption made in the pre-analysis that the influence of parameters other than net growth parameters will be
displayed if they are studied under a high net growth condition. The other important parameters are K_{navig} and K_s whose
influence is reduced by increased TOC. Moreover, very small interactions are observed between the parameters because almost

all of their global influence stem from their main effects ($S_T \approx S_1$ for each parameter), which once again confirms the previous
consideration that interactions are related to the effect of a varying net growth rate.

3.3 Third SA: Fixed low net growth

The results of the third SA (Fig. 8b) reveal that Y_{hb} is a predominantly influential parameter under a low net growth rate
condition across all TOC concentrations. This is due to the fact that the bacterial community hardly grows at all and therefore
BDOM is not a limiting factor for bacterial growth as there is not enough bacterial activity under a low net growth rate
condition. This result verifies the assumption made in the second SA by displaying all previous influential parameters except
Y_{hb} as non-influential.

In observing the role of BDOM, we conclude that by increasing the net growth rate, the influence of BDOM increases. Here,
we need to highlight that the actual upper limit of net growth is 0.11/h (Wang et al., 2018) compared to the current value of
0.05/h because the upper limit of $\mu_{max,hb}$ is reduced in Table 1 in order to prevent complete DO depletion of simulations that
occur due to the combination of extreme values in Sobol sampling. Therefore, it can be envisaged that BDOM may have a
greater influence under the actual range of net growth rate.

4   Discussion

In this section, we highlight answers to the proposed research questions and discuss what consequences do our results have on
water quality monitoring in urban areas and on data assimilation.

4.1   ~~The~~ Hierarchy of the most influential parameters during ~~a summer low-flow~~ low flow period

This study confirms that over a 45-day post-bloom summer low-flow period and whatever the TOC concentration~~(first SA)~~, the
bacterial net growth ~~rates~~ rate represented by $mort_{hb}$, $\mu_{max,hb}$ and Y_{hb} ~~are predominant and have the most influence on~~ control
the DO evolution. ~~On the other hand, since a significant portion of the total sensitivities of the bacterial parameters arise from
inter-parameter interactions, by fixing both the $mort_{hb}$ and $\mu_{max,hb}$ under the concept of a fixed~~ This is in accordance with
the findings of the study by Wang et al. (2018), which was conducted over a 4-day simulation period. However, the mentioned
bacteria physiological parameters are not sufficient to describe DO variation because the OM repartition parameter, b_1, and the
OM degradation parameter, K_s, are also influential at low flow.

Conducting the Sobol' SA for 5-days under a high bacterial net growth rate ~~once at a maximum value to prohibit their
interactions with other parameters (second SA) and by studying the river over a shorter period of 5 days,~~ condition demonstrated
the significant influence of BDOM and Y_{hb} ~~appear to be influential parameters as well. It highlights the importance of a
proper characterisation of BDOM during summer low flowperiods, when heterotrophic respiration is at its highest level
due to a high water temperature. This illustrates that BDOM is~~ at low flow. This is in connection with the findings of
Hullar et al. (2006); Crump et al. (2003) that emphasized the importance of BDOM on bacteria population. This result is also
in accordance with Bailey and Ahmadi (2014) who consider the model boundary condition as influential on DO but they do not

480 specify which portion of the OM neither do they conduct their study at low flow. Nevertheless, this work is the first quantitative recognition of the role of BDOM on DO evolution that illustrates BDOM as the most influential fraction of the total organic matter entering a river system from its boundary conditions.

5 ~~Discussion~~

~~The results of this study highlight the importance of having a better knowledge of i) organic matter outflow from boundary~~
485 ~~conditions during low flow, and ii) in particular of its biodegradable portion (BDOM), iii) in addition to bacterial physiology for water quality modeling at low flow. Therefore, these results can be expected to have the following consequences on data assimilation and water quality monitoring in urban areas:~~

4.1 ~~Consequences of the results on data assimilation (DA) strategy~~

~~Inclusion of BDOM and other influential parameters from this study in data assimilation tools is crucial for improving their~~
490 ~~performance at low flow. Despite its early development in freshwater sciences,~~

4.1 Limitations of the sensitivity analysis

This study is conducted under carbon source conditions where TOC is varied between 1-10 mgC/L. It is also assumed that a carbon sink condition never happens. Indeed, carbon depletion would preclude the possibility to quantify the influence of the ~~usage of assimilation techniques remain limited among the scientific community (Cho et al., 2020; ?). Since the 1970's,~~
495 ~~implementations of assimilation techniques switched from extended Kalman filter (EKF) (Beck and Young, 1976; Beck, 1978; Bowles and~~ ~~to ensemble Kalman filter (EnKF) (Beck and Halfon, 1991; Huang et al., 2013; Kim et al., 2014; Page et al., 2018; Zhang et al., 2020; Loo~~ ~~, which remains today the most popular assimilation method (Carrassi et al., 2018; Cho et al., 2020) that mostly focuses on phytoplankton dynamics in rivers and lakes or situations other than a low flow where DO is the state-variable.~~ carbon on the DO concentration. Therefore the upper limit of $\mu_{max,hb}$ is reduced from 0.13/h to 0.07/h so as to avoid depletion of both carbon
500 and oxygen during the simulation period.

On the other hand, ~~in the first application of a particle filter for data assimilation in limnology using ProSe-PA software to study a dry year (?Wang, 2019a; ?), mismatches were found between simulations and observation data in low-flow periods. This is due to the degeneracy effect of the low flow data assimilation tool which cannot reach oxygen concentrations as low as those observed in the natural environment and it is considered to be due to insufficient biodegradable organic matter~~
505 ~~loading in the model caused by underestimated BDOM inputs to the Seine river (?). Therefore, we propose the incorporation of BDOM (the most influential OM-related parameter through the b_1 parameter) as a new component of the particle filter implemented in ProSe-PA to facilitate its estimation and consequently improve the simulation results. Consistent with the results of the second and third SA, it would be appropriate to fix one or both of the $mort_{hb}$ and $\mu_{max,hb}$ parameters in order to decrease the parameter interactions and thereby better characterize the distribution of BDOM and other model parameters~~
510 ~~under low flow. In addition, to facilitate the inclusion of BDOM in any data assimilation tool~~ since C-RIVE currently lacks

the radiation effects of bacteria population, this process is not considered in this work. Nevertheless, it is recommended to be considered in future researches as radiation is found to damage the bacteria DNA impacting their growth and oxygen intake rates (Matallana-Surget and Wattiez, 2013). This phenomenon could be included in the model in a simplistic approach either by increasing bacterial mortality or decreasing its growth rate by a factor. However, it is necessary to ~~explicitly include the~~

515 ~~model of OM sharing (Fig.2) as part of the software, which will read TOC as an input data from now on.~~

~~Moreover, BDOM from each organic matter source (tributary river, WWTP, and CSO) should be independently represented in the DA scheme due to its distinct role and contribution of organic matter in urban rivers (Servais and Garnier, 1993; Seidl et al., 1998b). The fact that the effluents from WWTP and CSO contain large bacteria which have significantly different physiological properties than the small autochtonous ones (Garnier et al., 1992; Servais and Garnier, 1993) also advocate for the identification~~

520 ~~of input data following the three aforementioned pools. To represent the contrasted concentrations of bacteria upstream and downstream metropolis such as Paris urban area, two bacterial communities should be taken into account in the forward model. A strategy consisting of reducing the degree of freedom of the mathematical problem should also be implemented in the DA framework such as fixing a few parameters values of each community, for instance, their growth and mortality rates assuming that the net growth of small bacteria is slow while that of large bacteria is fast.~~

525 find experimental data in order to find this factor because the effects of radiation also depend on the type of bacterial community (Fernandes et al., 2021).

4.2 Consequences of the results on water quality monitoring in urban areas

~~The results of this study suggest the implementation of a densified (optimally designed) water quality monitoring system using high-frequency sensors to measure the pertinent information necessary for improving the model uncertainties and monitoring~~

530 ~~of the river metabolism~~ Rivers are highly sensitive to urban outflows at low flow ~~. In addition, these results propose regular monitoring of bacterial populations to detect any potential bacterial community shift following infrastructure developments in WWTP and transient events such as CSO. Therefore, continuous monitoring is essential for overall ecosystem management and for studies of human impacts (Beaulieu et al., 2013).~~

4.2.1 ~~How to densify the urban monitoring networks?~~

535 ~~Considering the importance of the contribution of organic matter from boundary conditions during low-flow periods, we propose the establishment of monitoring stations within a denser monitoring network capable of characterizing the upstream tributary rivers and the outflow of anthropogenic sources~~ (Seidl et al., 1998a; Huang et al., 2017) due to their low dilution capacity at this period of the year. Moreover, the construction or renewal of sanitary facilities, such as WWTP and CSO ~~because an appropriate spatial resolution (Polus et al., 2010) and consideration of point and non-point pollution sources (Dixon et al., 1999; Do et al~~

540 ~~is necessary in the design of monitoring networks. Ouyang (2005) used principal component analysis (PCA) to identify the necessary locations for establishing a new station or removal of previous stations in addition to specifying DOM and TOC among the main important variables that should be monitored.~~

~~This new constraint on the development and densification of monitoring networks is in addition to the classically proposed pragmatic criteria such as the (i) existence of pollution sources and variability~~ during transient events, induce changes in

545 water quality ~~between the existing monitoring stations, (ii) infrastructure for ease of accessibility to the station, and (iii) the location of tributaries to the system (Chilundo et al., 2008; Anvari et al., 2009) and less frequently the spatial variability of processes (Polus et al., 2010). Meanwhile, Strobl et al. (2006) have considered the information on land use such as point and non-point outflow sources and the watershed soil characteristics, hydrology, and topography in order to select the monitoring points; generally, there is no correlation between the size and the number of monitoring stations (Nguyen et al., 2019). Another~~

550 ~~practical way is to observe the discontinuities or jumps in DO profiles obtained from modeling software, which may help in determining the location of monitoring stations.~~

4.2.1 ~~How important are high-frequency data in water quality modeling?~~

~~We propose the collection of high-frequency BDOM data from boundary conditions as being indispensable in improving the model uncertainties and detecting the concentration peaks. Indeed, the hydro-biogeochemical processes take place on~~

555 ~~different time scales ranging from minutes and hours to weeks and months (Tomlinson and Carlo, 2003; Vilmin et al., 2018) . In order to be able to study both the trajectories of the systems (long term) (Flipo et al., 2021) and the transient events such as storm overflows, it is necessary to maintain high frequency monitoring networks over time. Such networks allow us to follow the influence of WWTP discharges and also the influence of storm overflows which provoke longitudinal decreases of downstream O_2 due to the direct relation between bacterial respiration and the increase of the BDOM (Stanley et al., 2012).~~

560 ~~Therefore, the conventional sampling campaigns followed by incubation experiments (Servais et al., 1995) for determining BDOM cannot meet the need for models and water managers. Thus, the sampling frequency should be set according to the~~ due to changes in DOM, and BDOM concentrations (Servais and Garnier, 1993; Seidl et al., 1998b). As a result, this induces potential shifts in bacterial communities, which have been found to be related to DOC source and its biodegradability (Hullar et al., 2006; Crump et al., 2003) such that an increase in BDOM is considered to increase the diversity of bacterial

565 populations (Landa et al., 2013). Even et al. (2004) therefore recommend a regular reassessment of the influential bacterial parameters. The sampling frequency in the monitoring stations should be set-up not only considering the temporal variability of the ~~desired variable~~ variable of interest, but also integrating possible successions of species.

~~However, the operational and economic aspects of any monitoring system should be optimized so as to keep the costs at a minimum while meeting the monitoring requirements (Dixon and Chiswell, 1996; Camara et al., 2020). Therefore, optimal or~~

570 ~~minimum sampling frequencies have been suggested for different water quality variables depending on the presence of outflows and their corresponding contribution to the river, the degree of acceptable uncertainty in the model, the monitoring goals, and the variability of the target variable itself (Chappell et al., 2017; Vilmin et al., 2018)~~ On the one hand, our study provides the list of important parameters for water quality modeling at low flow. On the other hand, these parameters are dynamic (varying with time). Thus in urban areas submitted to large WWTP outflows, continuous monitoring networks would be able to provide

575 up-to-date values of parameters that influence the water quality.

4.2.1 ~~Monitoring of bacterial communities is necessary for detection of community shift~~

Considering the results of this study ~~, which introduce~~ which introduce initially several bacterial parameters (mortality ~~and growth rates,~~ rate, maximum growth rate and growth yield) as having the most influence on DO evolution, we propose the implementation of regular bacterial community sampling campaigns in ~~order to provide updated information on the physiological properties of bacterial communities in the case of an autochtonous community shift and modifications of waste water treatment processes~~ the framework of the water quality monitoring program. Samples should thus be taken in river waters upstream from the major urban areas, and in the effluents of major WWTPs and CSOs. ~~This issue arises because, on the one hand, rivers are highly sensitive to urban outflows~~

Then, considering the importance of BDOM at low flow ~~(Seidl et al., 1998a; Huang et al., 2017) due to their low dilution capacity and, on the other hand, because with the construction of a new WWTP or any infrastructure development inside a WWTP as well as following transient events such as overflow from a CSO , the outflow quality changes in terms of bacteria, DOM, and~~ periods, we propose the addition of BDOM measurement in the existing and new monitoring stations. These stations shall be dense and capable of characterizing the upstream tributary rivers and the outflow of anthropogenic sources such as WWTP and CSO because an appropriate spatial resolution (Polus et al., 2010) and consideration of point and non-point pollution sources (Dixon et al., 1999; Do et al., 2012) are necessary to cover all BDOM sources. On the other hand, innovative methods such as specific ultraviolet absorbance at 254 nm ($SUVA_{254}$) and fluorescence spectroscopy (Parlanti et al., 2000, 2002; Goffin et could be proposed for high frequency measurement of OM and BDOM specifically.

4.3 Consequences of the results on data assimilation (DA) in water quality modeling

Coupling water quality models with data assimilation techniques allows to avoid relying on a continuous water quality monitoring. Data assimilation is a method that combines observation data and physically based modeling in a statistical framework. It consists in sequentially updating the model parameters so that the output of the model will match the observation at each time step (Carrassi et al., 2018). These techniques not only estimate the evolution of influential parameters thanks to observation data such as that of DO, but also provide enhanced simulation results of state variables.

The first assimilation tool that uses the particle filter technique (a statistical technique where numerous simulations are launched instead of one single simulation and are weighted based on how well they reproduce the observed data) to couple with a water quality model is the ProSe-PA software (Wang et al., 2019b; Wang, 2019a; Wang et al., 2022). While studying a dry year using this software, mismatches were found between simulated and observed DO in low-flow periods (Wang et al., 2022). These mismatches were assumed to be due to insufficient biodegradable organic matter loading in the model caused by underestimated BDOM inputs to the Seine river. Therefore, based on our SA results, we can confirm that hypothesis and propose the incorporation of BDOM (the most influential OM-related parameter through the b_1 parameter) as a new component of ProSe-PA. It will facilitate not only the estimation of BDOM~~concentrations (Servais and Garnier, 1993; Seidl et al., 1998b) . As a result, this induces potential shifts in the bacterial community, which have been found to be related to DOC source and its biodegradability (Hullar et al., 2006; Crump et al., 2003) such that an increase in BDOM is considered to increase~~

~~the diversity of bacterial populations (Landa et al., 2013); therefore, one could recommend the need for regular reassessment~~
~~of the influential bacterial parameters (Even et al., 2004). ?Wang (2019a); ? have also considered the lack of information on~~
~~heterotrophic bacterial communities as another source of mismatch in modeling at low flow. Indeed, here we can propose that~~
~~the sampling frequency in the monitoring stations must now be considered not only according to the temporal variability of~~
~~the variable of interest~~, but also ~~according to the possible successions of species. Moreover, such updated data on bacterial~~
~~communities would also help in narrowing down the variation ranges of the parameters, which would improve water quality~~
~~model performances, especially of data assimilation softwares~~ consequently improves the simulation results.

Another consequence of our results is that in accordance with the results of the second and third SA, it would be appropriate to fix one or both of the $mort_{hb}$ and $\mu_{max,hb}$ parameters during data assimilation in order to facilitate parameter identification. In addition, the model of OM repartition (Fig. 2) shall be explicitly included in the data assimilation software and BDOM from each organic matter source (tributary river, WWTP, and CSO) should be independently represented in the DA scheme due to their distinct contribution of organic matter and heterotrophic bacteria in urban rivers (Garnier et al., 1992; Servais and Garnier, 1993; Seidl .

Finally, we ~~believe~~ can conclude that once a water quality model capable of data assimilation is validated and its uncertainties are suitably reduced, it can provide acceptable estimates of water quality variables in periods when monitoring is not possible or at locations where accessibility is an issue (Reis et al., 2015; Jiang et al., 2020). ~~The coupling between water~~
~~quality monitoring networks and particle filter modeling softwares will lead to a very powerful generation of water quality~~
~~spatio-temporal interpolators that will be able to estimate water quality state indicators with a unpreceeded resolution and~~
~~accuracy.~~ In other words, ~~if the approach followed by~~ while the approach of PIREN-Seine (https://www.piren-seine.fr/en) research program ~~since its beginnings has shown all its power (Carré et al., 2021) via the determination of the properties of~~
~~the species communities by~~ followed laboratory incubation experiments (Servais et al., 1995), it appears today that the coupling between measurement networks and modeling makes it possible to approach the functioning of the systems in a more dynamic way ~~by the identification of the communities parameters that vary with time (?). This would lay the foundation for~~
~~using the model in forecast studies (Park et al., 2020) and for assessing different river management strategies (Reis et al., 2015)~~
~~.~~ (Wang et al., 2022).

5 Conclusions

The objective of this work was to investigate the role of organic matter loadings to river systems and the physiology of bacteria in river metabolism during a summer non-bloom low-flow period. New parameters were introduced to account for repartition and degradation of OM. Then, the sensitivity of the C-RIVE model was analyzed against the newly introduced and the already existing model parameters. The following conclusions can be drawn from this study:

– ~~First, the~~ The Sobol sensitivity analysis method proved very efficient in the identification of influential parameters on DO evolution in the C-RIVE model. Then, by fixing the interaction-inducing parameters, the influence of other parameters

was assessed~~;~~. This methodology may also be of interest for future sensitivity analysis where parameter interactions may hide the effect of other parameters;

- The net growth rate of bacteria composed of their maximum growth rate ($\mu_{max,hb}$) and mortality rate ($mort_{hb}$) is the most important parameter under different TOC concentrations; therefore, it is essential to have a better estimation of the variation ranges of the growth and mortality rates of bacterial communities;

- Model response is very sensitive to the biodegradable share of DOM (BDOM) contributed by the boundary conditions. The effect of this parameter prevails at higher net growth rates occurring during summer low-flow periods when the organic matter attributable to human pressure is abundant in the river;

- The river metabolism is dominated by bacterial activity at low flow during summer non-bloom periods;

- Water quality monitoring networks ~~should be densified to encompass all important BDOM contributing boundary conditions. BDOM data should be collected at high frequency~~shall continuously measure the influential parameters of this study in order to provide the water quality models with update values;

- More frequent sampling of autochtonous bacteria communities upstream and downstream of major urban areas and in majors WWTP and CSO effluents will be of considerable interest to validate time varying parameter values estimated by data assimilation frameworks;

- The results of this study provide ~~us with~~a list of influential and non-influential parameters. The latter can be fixed at their average or preferred value as per the literature, and the former can be introduced to the data assimilation tools in order to estimate their temporal evolution thanks to the high-frequency ~~data observed , such as DO .~~observed DO data;

Code availability. Name of software : C-RIVE

Contact address: nicolas.flipo@minesparis.psl.eu

Year first available: 2019

Program language: ANSI C

Operating system: Linux

Software access: Deposit in progress

Availability and cost: Open source

Licence: Eclipse Public Licence v2.0

Author contributions. The author contributions are as follows:

The calculations were designed by Masihullah Hasanyar, Thomas Romary and Shuaitao Wang;

The calculations were executed by Masihullah Hasanyar;

The results were analyzed by all authors;

The first draft of the article was drafted by Masihullah Hasanyar, and then it was revised and corrected by all authors;

And finally Nicolas Flipo did the search for funding;

Competing interests. The authors declare that there are no known conflicts of interest associated with this work and there has been no significant financial support for this study that could have influenced its outcome. The authors alone are responsible for the content and writing of this article.

Acknowledgements. This work is a contribution to the PIREN-SEINE research program, part of the french Long-Term Socio-Ecological Research (LTSER) site "Zones Ateliers Seine."

References

Anvari, A., Reyes, J. D., Esmaeilzadeh, E., Jarvandi, A., Langley, N., and Navia, K. R.: Designing an Automated Water Quality Monitoring System for West and Rhode Rivers, in: 2009 Systems and Information Engineering Design Symposium, pp. 131–136, https://doi.org/10.1109/SIEDS.2009.5166167, 2009.

Bailey, R. T. and Ahmadi, M.: Spatial and Temporal Variability of In-Stream Water Quality Parameter Influence on Dissolved Oxygen and Nitrate within a Regional Stream Network, Ecological Modelling, 277, 87–96, https://doi.org/10.1016/j.ecolmodel.2014.01.015, 2014.

Beaulieu, J. J., Arango, C. P., Balz, D. A., and Shuster, W. D.: Continuous Monitoring Reveals Multiple Controls on Ecosystem Metabolism in a Suburban Stream, Freshwater Biology, 58, 918–937, https://doi.org/10.1111/fwb.12097, 2013.

Beck, M.: Random signal analysis in an environmental sciences problem, Applied Mathematical Modelling, 2, 23–29, https://doi.org/10.1016/0307-904X(78)90032-X, 1978.

Beck, M. and Halfon, E.: Uncertainty, identifiability and the propagation of prediction errors: A case study of lake ontario, Journal of Forecasting, 10, 135–161, https://doi.org/10.1002/for.3980100109, 1991.

Beck, M. and Young, P.: Systematic identification of DO-BOD model structure, Jnl. of Env. Eng. Division , American Society of Civil Engineers, 102, 909–927, 1976.

Billen, G.: Protein Degradation in Aquatic Environments, in: Microbial Enzymes in Aquatic Environments, edited by Brock, T. D. and Chróst, R. J., pp. 123–143, Springer New York, New York, NY, https://doi.org/10.1007/978-1-4612-3090-8_7, 1991.

Billen, G., Lancelot, C., Becker, E., and Servais, P.: Modelling Microbial Processes (Phyto- and Bacterioplankton) in the Scheldt Estuary, Hydrobiological Bulletin, 22, 43–55, https://doi.org/10.1007/BF02256781, 1988.

Billen, G., Garnier, J., and Hanset', P.: Modelling Phytoplankton Development in Whole Drainage Networks: The RIVERSTRAHLER Model Applied to the Seine River System, Hydrobiologia, 289, 119–137, https://doi.org/10.1007/BF00007414, 1994.

Billen, G., Garnier, J., Ficht, A., and Cun, C.: Modeling the Response of Water Quality in the Seine River Estuary to Human Activity in Its Watershed over the Last 50 Years, Estuaries, 24, 977, https://doi.org/10.2307/1353011, 2001.

Boudreau, B. P.: Diagenetic Models and Their Implementation, vol. 134, Springer, 1997.

Bowles, D. and Grenney, W.: Steady state river quality modeling by sequential extended Kalman filters, Water Resources Research, 14, 84–96, https://doi.org/10.1029/WR014i001p00084, 1978.

Brookes, V. J., Jordan, D., Davis, S., Ward, M. P., and Heller, J.: Saltelli Global Sensitivity Analysis and Simulation Modelling to Identify Intervention Strategies to Reduce the Prevalence of Escherichia Coli O157 Contaminated Beef Carcasses, PLoS One, 10, 1–22, https://doi.org/10.1371/journal.pone.0146016, 2015.

Camara, M., Jamil, N. R., Abdullah, A. F. B., binti Hashim, R., and Aliyu, A. G.: Economic and Efficiency Based Optimisation of Water Quality Monitoring Network for Land Use Impact Assessment, Science of The Total Environment, 737, 139 800, https://doi.org/10.1016/j.scitotenv.2020.139800, 2020.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, Wiley Interdisciplinary Reviews: Climate Change, 9, e535, https://doi.org/10.1002/wcc.535, 2018.

Carré, C., Meybeck, M., Garnier, J., Chong, N., Deroubaix, J.-F., Flipo, N., Goutte, A., Le Pichon, C., Seguin, L., and Tournebize, J.: The Seine River Basin, vol. 90 of *Handbook of Environmental Chemistry*, chap. River Basin Visions: Tools and Approaches from Yesterday to Tomorrow, pp. 381–414, Springer, Cham, Switzerland, https://doi.org/10.1007/698_2019_438, 2021.

Chappell, N. A., Jones, T. D., and Tych, W.: Sampling Frequency for Water Quality Variables in Streams: Systems Analysis to Quantify Minimum Monitoring Rates, Water Research, 123, 49–57, https://doi.org/10.1016/j.watres.2017.06.047, 2017.

Chilundo, M., Kelderman, P., and O´keeffe, J. H.: Design of a Water Quality Monitoring Network for the Limpopo River Basin in Mozambique, Physics and Chemistry of the Earth, Parts A/B/C, 33, 655–665, https://doi.org/10.1016/j.pce.2008.06.055, 2008.

Cho, K. H., Pachepsky, Y., Ligaray, M., Kwon, Y., and Kim, K. H.: Data Assimilation in Surface Water Quality Modeling: A Review, Water Research, 186, 116 307, https://doi.org/10.1016/j.watres.2020.116307, 2020.

Cosby, B. and Hornberger, G.: Identification of photosynthesis-light models for aquatic systems I. Theory and simulations, Ecological Modelling, 23, 1–24, https://doi.org/10.1016/0304-3800(84)90116-9, 1984.

Cox, B.: A Review of Currently Available In-Stream Water-Quality Models and Their Applicability for Simulating Dissolved Oxygen in Lowland Rivers, The Science of The Total Environment, 314-316, 335–377, https://doi.org/10.1016/S0048-9697(03)00063-9, 2003.

Crump, B. C., Kling, G. W., Bahr, M., and Hobbie, J. E.: Bacterioplankton Community Shifts in an Arctic Lake Correlate with Seasonal Changes in Organic Matter Source, Applied and Environmental Microbiology, 69, 2253–2268, https://doi.org/10.1128/AEM.69.4.2253-2268.2003, 2003.

Demars, B., Thompson, J., and Manson, J.: Stream Metabolism and the Open Diel Oxygen Method: Principles, Practice, and Perspectives, Limnology and Oceanography: Methods, 13, 356–374, https://doi.org/10.1002/lom3.10030, 2015.

Dixon, W. and Chiswell, B.: Review of Aquatic Monitoring Program Design, Water Research, 30, 1935–1948, https://doi.org/10.1016/0043-1354(96)00087-5, 1996.

Dixon, W., Smyth, G. K., and Chiswell, B.: Optimized Selection of River Sampling Sites, Water Research, 33, 971–978, https://doi.org/10.1016/S0043-1354(98)00289-9, 1999.

Do, H. T., Lo, S.-L., Chiueh, P.-T., and Phan Thi, L. A.: Design of Sampling Locations for Mountainous River Monitoring, Environmental Modelling & Software, 27-28, 62–70, https://doi.org/10.1016/j.envsoft.2011.09.007, 2012.

Escoffier, N., Bensoussan, N., Vilmin, L., Flipo, N., Rocher, V., David, A., Métivier, F., and Groleau, A.: Estimating Ecosystem Metabolism from Continuous Multi-Sensor Measurements in the Seine River, Environ Sci Pollut Res, 25, 23 451–23 467, https://doi.org/10.1007/s11356-016-7096-0, 2018.

Even, S., Poulin, M., Garnier, J., Billen, G., Servais, P., Chesterikoff, A., and Coste, M.: River Ecosystem Modelling: Application of the PROSE Model to the Seine River (France), Hydrobiologia, 373, 27–45, https://doi.org/10.1023/A:1017045522336, 1998.

Even, S., Poulin, M., Mouchel, J.-M., Seidl, M., and Servais, P.: Modelling Oxygen Deficits in the Seine River Downstream of Combined Sewer Overflows, Ecological Modelling, 173, 177–196, https://doi.org/10.1016/j.ecolmodel.2003.08.019, 2004.

Even, S., Mouchel, J. M., Servais, P., Flipo, N., Poulin, M., Blanc, S., Chabanel, M., and Paffoni, C.: Modeling the impacts of Combined Sewer Overflows on the river Seine water quality, Science of The Total Environment, 375, 140–151, https://doi.org/10.1016/j.scitotenv.2006.12.007, 2007.

Fernandes, A., Oliveira, A., Soares, R., and Barata, P.: The Effects of Ionizing Radiation on Gut Microbiota, a Systematic Review, Nutrients, 13, 3025, https://doi.org/10.3390/nu13093025, 2021.

Flipo, N.: Modélisation intégrée des transferts d'azote dans les aquifères et les rivières : Application au bassin du Grand Morin, Ph.D. thesis, CIG, 2005.

Flipo, N., Even, S., Poulin, M., Tusseau-Vuillemin, M.-H., Ameziane, T., and Dauta, A.: Biogeochemical Modelling at the River Scale: Plankton and Periphyton Dynamics: Grand Morin Case Study, France, Ecological Modelling, 176, 333–347, https://doi.org/10.1016/j.ecolmodel.2004.01.012, 2004.

Flipo, N., Rabouille, C., Poulin, M., Even, S., Tusseau-Vuillemin, M.-H., and Lalande, M.: Primary Production in Headwater Streams of the Seine Basin: The Grand Morin River Case Study, Science of The Total Environment, 375, 98–109, https://doi.org/10.1016/j.scitotenv.2006.12.015, 2007.

755 Flipo, N., Lestel, L., Labadie, P., Meybeck, M., and Garnier, J.: The Seine River Basin, vol. 90 of *Handbook of Environmental Chemistry*, chap. Trajectories of the Seine River Basin, pp. 1–28, Springer, Cham, Switzerland, https://doi.org/10.1007/698_2019_437, 2021.

Garnier, J. and Benest, D.: Seasonal Coupling between Phyto- and Bacterioplankton in a Sand Pit Lake (Creteil Lake, France), Hydrobiologia, 207, 71–77, 1990.

Garnier, J., Billen, G., Servais, P., and Servais, P.: Physiological Characteristics and Ecological Role of Small and Large Sized Bacteria in a
760 Polluted River (Seine River, France), Arch. Hydrobiol. Beih., 37, 83–94, 1992.

Garnier, J., Billen, G., and Coste, M.: Seasonal Succession of Diatoms and Chlorophyceae in the Drainage Network of the Seine River: Observation and Modeling, Limnology and Oceanography, 40, 750–765, https://doi.org/10.4319/lo.1995.40.4.0750, 1995.

Garnier, J., Billen, G., and Palfner, L.: Understanding the Oxygen Budget and Related Ecological Processes in the River Mosel: The RIVER-STRAHLER Approach, in: Man and River Systems, edited by Garnier, J. and Mouchel, J.-M., pp. 151–166, Springer Netherlands, Dor-
765 drecht, https://doi.org/10.1007/978-94-017-2163-9_17, 1999.

Garnier, J., Némery, J., Billen, G., and Théry, S.: Nutrient Dynamics and Control of Eutrophication in the Marne River System: Modelling the Role of Exchangeable Phosphorus, Journal of Hydrology, 304, 397–412, https://doi.org/10.1016/j.jhydrol.2004.07.040, 2005.

Garnier, J., Marescaux, A., Guillon, S., Vilmin, L., Rocher, V., Billen, G., Thieu, V., Silvestre, M., Passy, P., Raimonet, M., Groleau, A., Théry, S., Tallec, G., and Flipo, N.: Ecological Functioning of the Seine River: From Long-Term Modelling Approaches to High-Frequency Data
770 Analysis, in: The Handbook of Environmental Chemistry, pp. 1–28, Springer Berlin Heidelberg, 2020.

Garvey, J. E., Whiles, M. R., and Streicher, D.: A Hierarchical Model for Oxygen Dynamics in Streams, Can. J. Fish. Aquat. Sci., 64, 1816–1827, https://doi.org/10.1139/f07-144, 2007.

Goffin, A., Guérin, S., Rocher, V., and Varrault, G.: 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, Tech. rep., PIREN-Seine phase VII,
775 2017.

Hasanyar, M., Flipo, N., Romary, T., Wang, S., and Yari, A.: Rôle de la matière organique dans le métabolisme des rivières à bas débit, PIREN-Seine phase 8 - Technical report, PIREN-Seine, 2020.

Hasanyar, M., Flipo, N., Romary, T., and Wang, S.: A review of water quality variables and parameters characterizing organic matter in world rivers., manuscript in preparation, available upon request from masihullah.hasanyar@mines-paristech.fr, 2021.

780 Herman, J. and Usher, W.: SALib: An Open-Source Python Library for Sensitivity Analysis, Journal of Open Source Software, 2, 97, https://doi.org/10.21105/joss.00097, 2017.

Huang, J., Gao, J., Liu, J., and Zhang, Y.: State and parameter update of a hydrodynamic-phytoplankton model using ensemble Kalman filter, Ecological Modelling, 263, 81–91, https://doi.org/10.1016/j.ecolmodel.2013.04.022, 2013.

Huang, J., Yin, H., Chapra, S., and Zhou, Q.: Modelling Dissolved Oxygen Depression in an Urban River in China, Water, 9, 520,
785 https://doi.org/10.3390/w9070520, 2017.

Hullar, M. A. J., Kaplan, L. A., and Stahl, D. A.: Recurring Seasonal Dynamics of Microbial Communities in Stream Habitats, Applied and Environmental Microbiology, 72, 713–722, https://doi.org/10.1128/AEM.72.1.713-722.2006, 2006.

Hutchins, M., Harding, G., Jarvie, H., Marsh, T., Bowes, M., and Loewenthal, M.: Intense Summer Floods May Induce Prolonged Increases in Benthic Respiration Rates of More than One Year Leading to Low River Dissolved Oxygen, Journal of Hydrology X, 8, 100 056, https://doi.org/10.1016/j.hydroa.2020.100056, 2020.

Jiang, J., Tang, S., Han, D., Fu, G., Solomatine, D., and Zheng, Y.: A Comprehensive Review on the Design and Optimization of Surface Water Quality Monitoring Networks, Environmental Modelling & Software, 132, 104 792, https://doi.org/10.1016/j.envsoft.2020.104792, 2020.

Jolliffe, I. T. and Cadima, J.: Principal Component Analysis: A Review and Recent Developments, Phil. Trans. R. Soc. A, 374, 20150 202, https://doi.org/10.1098/rsta.2015.0202, 2016.

Kim, K., Park, M., Min, J., Ryu, I., Kang, M., and Park, L. J.: Simulation of algal bloom dynamics in a river with the ensemble Kalman filter, Journal of Hydrology, 519, 2810–2821, https://doi.org/10.1016/j.jhydrol.2014.09.073, 2014.

Landa, M., Cottrell, M. T., Kirchman, D. L., Blain, S., and Obernosterer, I.: Changes in Bacterial Diversity in Response to Dissolved Organic Matter Supply in a Continuous Culture Experiment, Aquatic Microbial Ecology, 69, 157–168, https://doi.org/10.3354/ame01632, 2013.

Larsson, U. and Hagstrom, A.: Phytoplankton Exudate Release as an Energy Source for the Growth of Pelagic Bacteria, Marine Biology, 52, 199–206, https://doi.org/10.1007/BF00398133, 1979.

Loos, S., Shin, C. M., Sumihar, J., Kim, K., Cho, J., and Weerts, A. H.: Ensemble data assimilation methods for improving river water quality forecasting accuracy, Water Research, 171, 115 343, https://doi.org/10.1016/j.watres.2019.115343, 2020.

Mao, J., Lee, J. H., and Choi, K.: The extended Kalman filter for forecast of algal bloom dynamics, Water Research, 43, 4214 – 4224, https://doi.org/10.1016/j.watres.2009.06.012, 2009.

Marescaux, A., Thieu, V., Gypens, N., Silvestre, M., and Garnier, J.: Modeling Inorganic Carbon Dynamics in the Seine River Continuum in France, Hydrology and Earth System Sciences, 24, 2379–2398, https://doi.org/10.5194/hess-24-2379-2020, 2020.

Martin, L.: Ecological functioning of the Seine river downstream the wastewater treatment plant "Seine Aval": experimental datas and bidimensional modeling., Ph.D. thesis, École Nationale Supérieure des Mines de Paris, Sciences of the Universe [physics], 2001.

Matallana-Surget, S. and Wattiez, R.: Impact of Solar Radiation on Gene Expression in Bacteria, Proteomes, 1, 70–86, https://doi.org/10.3390/proteomes1020070, 2013.

Nguyen, T. H., Helm, B., Hettiarachchi, H., Caucci, S., and Krebs, P.: The Selection of Design Methods for River Water Quality Monitoring Networks: A Review, Environmental Earth Sciences, 78, 96, https://doi.org/10.1007/s12665-019-8110-x, 2019.

Nossent, J., Elsen, P., and Bauwens, W.: Sobol' Sensitivity Analysis of a Complex Environmental Model, Environmental Modelling & Software, 26, 1515–1525, https://doi.org/10.1016/j.envsoft.2011.08.010, 2011.

Odum, H. T.: Primary Production in Flowing Waters1, Limnol. Oceanogr., 1, 102–117, https://doi.org/10.4319/lo.1956.1.2.0102, 1956.

Ouyang, Y.: Evaluation of River Water Quality Monitoring Stations by Principal Component Analysis, Water Research, 39, 2621–2635, https://doi.org/10.1016/j.watres.2005.04.024, 2005.

Page, T., Smith, P. J., Beven, K. J., Jones, I. D., Elliott, J. A., Maberly, S. C., Mackay, E. B., Ville, M. D., and Feuchtmayr, H.: Adaptive forecasting of phytoplankton communities, Water Research, 134, 74 – 85, https://doi.org/10.1016/j.watres.2018.01.046, 2018.

Park, S., Kim, K., Shin, C., Min, J.-H., Na, E. H., and Park, L. J.: Variable Update Strategy to Improve Water Quality Forecast Accuracy in Multivariate Data Assimilation Using the Ensemble Kalman Filter, Water Research, 176, 115 711, https://doi.org/10.1016/j.watres.2020.115711, 2020.

Parlanti, E., Wörz, K., Geoffroy, L., and Lamotte, M.: Dissolved Organic Matter Fluorescence Spectroscopy as a Tool to Estimate Biological

825   Activity in a Coastal Zone Submitted to Anthropogenic Inputs, Organic Geochemistry, 31, 1765–1781, https://doi.org/10.1016/S0146-6380(00)00124-8, 2000.

Parlanti, E., Morin, B., and Vacher, L.: Combined 3D-Spectrofluorometry, High Performance Liquid Chromatography and Capillary Electrophoresis for the Characterization of Dissolved Organic Matter in Natural Waters, Organic Geochemistry, 33, 221–236, https://doi.org/10.1016/S0146-6380(01)00154-1, 2002.

830   Pastres, R., Ciavatta, S., and Solidoro, C.: The Extended Kalman Filter (EKF) as a tool for the assimilation of high frequency water quality data, Ecological Modelling, 170, 227–235, https://doi.org/10.1016/S0304-3800(03)00230-8, 2003.

Polus, E., de Fouquet, C., Flipo, N., and Poulin, M.: Caractérisation spatiale et temporelle de la qualité des « Masses d'Eau Cours d'Eau », rseau, 23, 415–429, https://doi.org/10.7202/045101ar, 2010.

Polus, E., Flipo, N., de Fouquet, C., and Poulin, M.: Geostatistics for Assessing the Efficiency of a Distributed Physically-Based Water

835   Quality Model: Application to Nitrate in the Seine River, Hydrol. Process., 25, 217–233, https://doi.org/10.1002/hyp.7838, 2011.

Pryet, A., Labarthe, B., Saleh, F., Akopian, M., and Flipo, N.: Reporting of stream-aquifer flow distribution at the regional scale with a distributed process-based model, Water Resources Management, 29, 139–159, https://doi.org/10.1007/s11269-014-0832-7, 2015.

Raimonet, M., Vilmin, L., Flipo, N., Rocher, V., and Laverman, A.: Modelling the fate of nitrite in an urbanized river using experimentally obtained nitrifier growth parameters, Water Research, 73, 373–387, https://doi.org/10.1016/j.watres.2015.01.026, 2015.

840   Reis, S., Seto, E., Northcross, A., Quinn, N. W. T., Convertino, M., Jones, R. L., Maier, H. R., Schlink, U., Steinle, S., Vieno, M., and Wimberly, M. C.: Integrating Modelling and Smart Sensors for Environmental and Human Health, Environmental Modelling & Software, 74, 238–246, https://doi.org/10.1016/j.envsoft.2015.06.003, 2015.

Saltelli, A.: Making Best Use of Model Evaluations to Compute Sensitivity Indices, Computer Physics Communications, 145, 280–297, https://doi.org/10.1016/S0010-4655(02)00280-1, 2002.

845   Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S.: Variance Based Sensitivity Analysis of Model Output. Design and Estimator for the Total Sensitivity Index, Computer Physics Communications, 181, 259–270, https://doi.org/10.1016/j.cpc.2009.09.018, 2010.

Seidl, M., Servais, P., Martaud, M., Gandouin, C., and Mouchel, J. M.: Organic Carbon Biodegradability and Heterotrophic Bacteria along a Combined Sewer Catchment during Rain Events, Water Science and Technology, 37, 25–33, https://doi.org/10.1016/S0273-

850   1223(97)00752-X, 1998a.

Seidl, M., Servais, P., and Mouchel, J.: Organic Matter Transport and Degradation in the River Seine (France) after a Combined Sewer Overflow, Water Research, 32, 3569–3580, https://doi.org/10.1016/S0043-1354(98)00169-9, 1998b.

Servais, P.: Modélisation de la biomasse et de l'activité bactérienne dans la Meuse belge, rseau, 2, 543–563, https://doi.org/10.7202/705042ar, 1989.

855   Servais, P. and Garnier, J.: Contribution of Heterotrophic Bacterial Production to the Carbon Budget of the River Seine (France), Microb Ecol, 25, 19–33, https://doi.org/10.1007/BF00182127, 1993.

Servais, P., Barillier, A., and Garnier, J.: Determination of the Biodegradable Fraction of Dissolved and Particulate Organic Carbon in Waters, Ann. Limnol. - Int. J. Lim., 31, 75–80, https://doi.org/10.1051/limn/1995005, 1995.

Sobol, I. M.: Sensitivity Estimates for Nonlinear Mathematical Models, Mathematical modelling and computational experiments, 1, 407–414,

860   1993.

Staehr, P., Bade, D., van de Bogert, M., Koch, G., Williamson, C., Hanson, P., Cole, J., and Kratz, T.: Lake Metabolism and the Diel Oxygen Technique: State of the Science, Limnology and Oceanography: Methods, 8, 628–644, https://doi.org/10.4319/lom.2010.8.0628, 2010.

Stanley, E. H., Powers, S. M., Lottig, N. R., Buffam, I., and Crawford, J. T.: Contemporary changes in dissolved organic carbon (DOC) in human-dominated rivers: is there a role for DOC management?, Freshwater Biology, 57, 26–42, https://doi.org/https://doi.org/10.1111/j.1365-2427.2011.02613.x, 2012.

Streeter, H. and Phelps, E. B.: A study of the pollution and natural purification of the Ohio River, Tech. Rep. 146, U.S. Public Health Service, Treasury Department, Washington DC, 1925.

Strobl, R. O., Robillard, P. D., Day, R. L., Shannon, R. D., and McDonnell, A. J.: A Water Quality Monitoring Network Design Methodology for the Selection of Critical Sampling Points: Part II, Environmental Monitoring and Assessment, 122, 319–334, https://doi.org/10.1007/s10661-006-0358-4, 2006.

Thibodeaux, L., Poulin, M., and Even, S.: A Model for Enhanced Aeration of Streams by Motor Vessels with Application to the Seine River, Journal of Hazardous Materials, 37, 459–473, https://doi.org/10.1016/0304-3894(93)E0101-7, 1994.

Thieu, V., Billen, G., and Garnier, J.: Nutrient Transfer in Three Contrasting NW European Watersheds: The Seine, Somme, and Scheldt Rivers. A Comparative Application of the Seneque/Riverstrahler Model, Water Research, 43, 1740–1754, https://doi.org/10.1016/j.watres.2009.01.014, 2009.

Tomlinson, M. S. and Carlo, E. H. D.: The Need for High Resolution Time Series Data to Characterize Hawaiian Streams1, JAWRA Journal of the American Water Resources Association, 39, 113–123, https://doi.org/10.1111/j.1752-1688.2003.tb01565.x, 2003.

Vilmin, L.: Modélisation Du Fonctionnement Biogéochimique de La Seine de l'agglomération Parisienne à l'estuaire à Différentes Échelles Temporelles, These de doctorat, Paris, ENMP, 2014.

Vilmin, L., Flipo, N., and Poulin, M.: Le modèle de simulation biogéochimique C-RIVE, Tech. rep., PIREN-Seine, 2012.

Vilmin, L., Aissa-Grouz, N., Garnier, J., Billen, G., Mouchel, J. M., Poulin, M., and Flipo, N.: Impact of hydro-sedimentary processes on the dynamics of soluble reactive phosphorus in the Seine River, Biogeochemistry, 122, 229–251, https://doi.org/10.1007/s10533-014-0038-3, 2015a.

Vilmin, L., Flipo, N., de Fouquet, C., and Poulin, M.: Pluri-Annual Sediment Budget in a Navigated River System: The Seine River (France), Science of The Total Environment, 502, 48–59, https://doi.org/10.1016/j.scitotenv.2014.08.110, 2015b.

Vilmin, L., Flipo, N., Escoffier, N., Rocher, V., and Groleau, A.: Carbon Fate in a Large Temperate Human-Impacted River System: Focus on Benthic Dynamics, Global Biogeochem. Cycles, 30, 1086–1104, https://doi.org/10.1002/2015GB005271, 2016.

Vilmin, L., Flipo, N., Escoffier, N., and Groleau, A.: Estimation of the water quality of a large urbanized river as defined by the European WFD: what is the optimal sampling frequency?, ESPR, 25, 23 485–23 501, https://doi.org/10.1007/s11356-016-7109-z, 2018.

Wang, Q., Li, S., Jia, P., Qi, C., and Ding, F.: A Review of Surface Water Quality Models, The Scientific World Journal, 2013, e231 768, https://doi.org/10.1155/2013/231768, 2013.

Wang, S.: Simulation Du Métabolisme de La Seine Par Assimilation de Données En Continu, These de doctorat, Paris Sciences et Lettres, 2019a.

Wang, S., Flipo, N., and Romary, T.: Time-Dependent Global Sensitivity Analysis of the C-RIVE Biogeochemical Model in Contrasted Hydrological and Trophic Contexts, Water Research, 144, 341–355, https://doi.org/10.1016/j.watres.2018.07.033, 2018.

Wang, S., Flipo, N., and Romary, T.: Oxygen Data Assimilation for Estimating Micro-Organism Communities' Parameters in River Systems, Water Research, 165, 115 021, https://doi.org/10.1016/j.watres.2019.115021, 2019b.

Wang, S., Flipo, N., Romary, T., and Hasanyar, M.: Particle Filter for High Frequency Oxygen Data Assimilation in River Systems, Environmental Modelling & Software, 151, 105 382, https://doi.org/10.1016/j.envsoft.2022.105382, 2022.

900   Whitehead, P. and Hornberger, G.: Modelling algal behaviour in the river thames, Water Research, 18, 945–953, https://doi.org/10.1016/0043-1354(84)90244-6, 1984.

Wold, S., Esbensen, K., and Geladi, P.: Principal Component Analysis, Chemometrics and Intelligent Laboratory Systems, 2, 37–52, 1987.

Yang, C.-P., Lung, W.-S., Kuo, J.-T., and Liu, J.-H.: Water Quality Modeling of a Hypoxic Stream, Pract. Period. Hazard. Toxic Radioact. Waste Manage., 14, 115–123, https://doi.org/10.1061/(ASCE)HZ.1944-8376.0000021, 2010.

905   Zhang, C., Yan, Q., Kuczyńska-Kippen, N., and Gao, X.: An Ensemble Kalman Filter Approach to Assess the Effects of Hydrological Variability, Water Diversion, and Meteorological Forcing on the Total Phosphorus Concentration in a Shallow Reservoir, Science of The Total Environment, 724, 138 215, https://doi.org/10.1016/j.scitotenv.2020.138215, 2020.