Supplement of Biogeosciences, 20, 1621–1633, 2023
https://doi.org/10.5194/bg-20-1621-2023-supplement

Biogeosciences Open Access EGU

Supplement of

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

Masihullah Hasanyar et al.

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

S1 DO and OM processes in C-RIVE

S1.1 Dissolved oxygen evolution equations

DO in the water column depends on physical, bacterial, and phytoplanktonic processes:

$$\frac{d[O_2]}{dt} = \frac{d[O_2]}{dt}_{physical} + \frac{d[O_2]}{dt}_{phytoplanktonic} + \frac{d[O_2]}{dt}_{bacterial} \tag{S1}$$

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:

$$\frac{d[O_2]}{dt}_{physical} = \frac{K_{rea}}{h}([O_2]_{sat}(T) - [O_2]) - \frac{D_s}{h}([O_2]_{water} - [O_2]_{sed}) + \frac{d[O_2]}{dt}_{dams} \tag{S2}$$

where,

h: water depth [m]

$[O_2]_{sat}(T)$: the saturated oxygen concentration at temperature T [mgO_2L^{-1}]

D_s: the coefficient of diffusion between water and sediment layer [ms^{-1}]

K_{rea} : [ms^{-1}] the reoxygenation coefficient calculated from the empirical formula of Thibodeaux et al. (1994) as follows:

$$K_{rea} = \sqrt{\frac{D_m V_{wat}}{h}} + (K_{wind} V_{wind}^{2.23} (D_m * 10^4)^{2/3} + K_{navig}) \tag{S3}$$

where,

K_{wind}: reoxygenation coefficient due to wind [ms^{-1}]

V_{wind}: wind speed [ms^{-1}]

V_{wat}: river flow velocity [ms^{-1}]

K_{navig}: reoxygenation coefficient due to navigation [ms^{-1}] (Vilmin, 2014)

D_m: molecular diffusivity of DO [m^{2}s^{-1}]

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

$$\frac{d[O_2]}{dt}_{phytoplanktonic} = P_{O_2,pp} - R_{O_2,pp} \tag{S4}$$

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:

$$\frac{d[O_2]}{dt}_{bacterial} = -\tau_{HB}(1 - Y_{HB})\ \textit{uptake} \tag{S5}$$

$$\textit{uptake} = \frac{\mu_{max,HB} e^{\frac{-(T-T_{opt,HB})^2}{\sigma_{HB}^2}} \frac{[S]}{[S]+K_S} [HB]}{Y_{HB}} \tag{S6}$$

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) [$mgCL^{-1}$]

S1.2 Organic matter degradation equations

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. (S7) and (S9). 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 (Larsson and Hagstrom, 1979; Garnier and Benest, 1990; Billen, 1991).

$$\frac{d[S = DOM_1]}{dt} = hyd_{DOM_2} + hyd_{POM_{1,2}} - uptake_{HB} + P_S + P_E + P_L \tag{S7}$$

where,
P_S, P_E, P_L: represent DOM_1 from the direct input of OM sources, phytoplankton excretion, and microorganism lysis, respectively [$mgCL^{-1}h^{-1}$]
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 [$mgCL^{-1}h^{-1}$]

$$hyd_{DOM_2} = k_{hyd,max}\frac{[DOM_2]}{[DOM_2] + K_{DOM2}}[HB] \tag{S8}$$

$$uptake_{HB} = \mu_{max,HB}\frac{[DOM_1]}{[DOM_1] + K_s}[HB] \tag{S9}$$

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 [$mgCL^{-1}$]

S1.3 Parameterization of organic matter partitioning and degradation

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

S1.3.1 OM degradation parameters

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 S1.2 and that already exist in C-RIVE. 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).

S1.3.2 OM partitioning parameters

The following five parameters are introduced to represent the partitioning of OM:

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

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

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

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

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

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 [-]
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 [-]

$DOM_{1,2,3}$ and $POM_{1,2,3}$ were state variables in the former version of C-RIVE. They used to be forced information provided by user. They are now defined by the proposed partitioning model which has the above-mentioned five parameters.

S2 SA methodology steps

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 analysis 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 S1 demonstrates the ensemble of 260,000 [=N(2D+2)] DO simulations of TOC = 5 $mgCL^{-1}$ in the second SA analysis (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

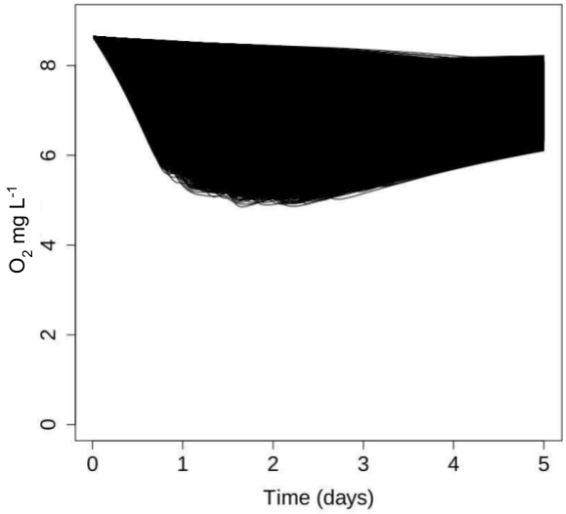

Figure S1. Ensemble of the 260,000 DO simulations for TOC = 5 mgCL^{-1} in the second SA analysis

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 (principal 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. S2a), where four (k) significant EOFs are found such that the first EOF (EOF_1) represents almost 55% of the total variance. Figure S2b illustrates the evolution of the eigenvalues of the four (k) EOFs with time, which are 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.

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:

$$Y = f(X) = f(X_1, ..., X_D), \tag{S10}$$

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

$$f(X) = f_0 + \sum_{i=1}^{D} f_i(X_i) + \sum_{i=1}^{D-1} \sum_{j=i+1}^{D} f_{ij}(X_i, X_j) + ... + f_{1,...,D}(X_1, ..., X_D) \tag{S11}$$

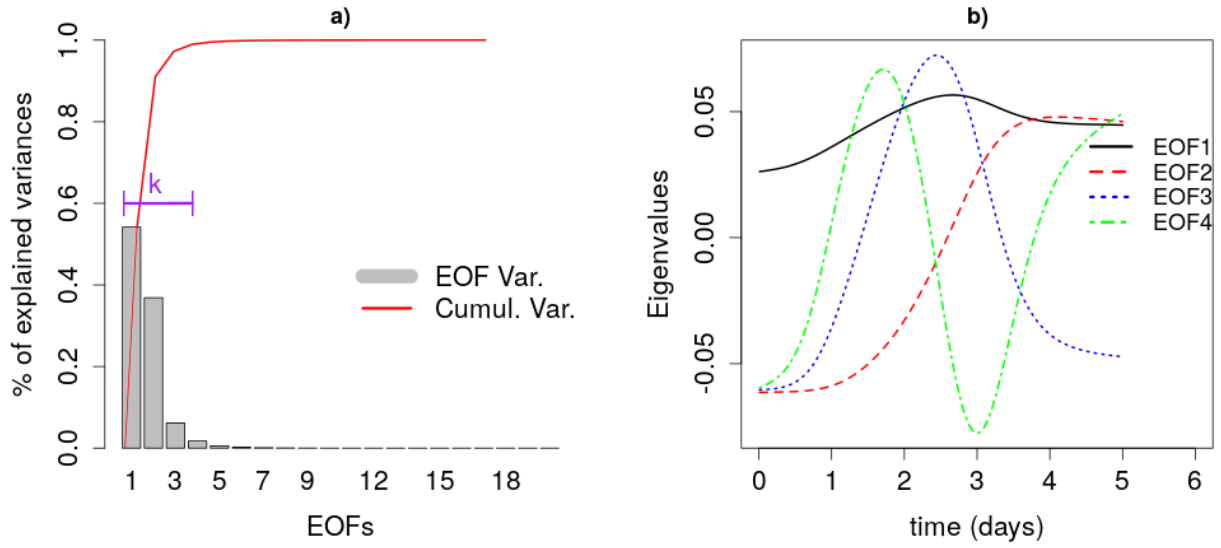

Figure S2. a) Cumulative sum of EOF variances and b) time evolution of four (k) significant EOFs for TOC = 5 mgCL^{-1} in second SA

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:

$$f_0 = \int_{[0,1]^D} f(X) d_X \tag{S12}$$

$$f_i(X_i) = -f_0 + \int_{[0,1]^{D-1}} f(X) d_{X \sim i} \tag{S13}$$

$$f_{ij}(X_i, X_j) = -f_0 - f_i(X_i) - f_j(X_j) + \int_{[0,1]^{D-2}} f(X) d_{X \sim (ij)} \tag{S14}$$

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

$$V(Y) = \int_{[0,1]^D} f^2(X) d_X - f_0^2 \tag{S15}$$

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

$$V(Y) = \sum_{i=1}^{D} V_i(X_i) + \sum_{i=1}^{D-1} \sum_{j=i+1}^{D} V_{ij}(X_i, X_j) + ... + V_{1,...,D}(X_1, ..., X_D) \tag{S16}$$

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:

$$V_{i_1,...,i_s} = \int_0^1 ... \int_0^1 f^2_{i_1,...,i_s}(X_{i_1}, ..., X_{i_s}) dX_{i_1}, ...dX_{i_s} \tag{S17}$$

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

$$S_i = \frac{V_i}{V} \tag{S18}$$

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:

$$S_{Ti} = S_i + \sum_{j \neq i} S_{ij} + ... \tag{S19}$$

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:

$$S_{2,A} = \sum_{j} S_{Aj} \tag{S20}$$

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.

References

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 Schelde 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.

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.

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

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.

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.

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.

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.

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.

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.

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.

Sobol, I. M.: Sensitivity Estimates for Nonlinear Mathematical Models, Mathematical modelling and computational experiments, 1, 407–414, 1993.

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.

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.

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