



Global sensitivity analysis in the development of first principle-based eutrophication models

V. Estrada^{a,b}, M.S. Diaz^{a,c,*}

^a PLAPIQUI, Planta Piloto de Ingeniería Química (UNS – CONICET), Camino La Carrindanga Km 7, Bahía Blanca, Argentina

^b Department of Biology, Biochemistry and Pharmacy, Universidad Nacional del Sur, Alem 1253, Bahía Blanca, Argentina

^c Department of Chemical Engineering, Universidad Nacional del Sur, Alem 1253, Bahía Blanca, Argentina

ARTICLE INFO

Article history:

Received 25 September 2009

Received in revised form

27 June 2010

Accepted 28 June 2010

Available online 4 August 2010

Keywords:

Global sensitivity analysis

Sensitivity indices

First principle-based water quality model

Eutrophication

ABSTRACT

In this work, we formulate a dynamic first principle-based eutrophication model for a reservoir and perform global sensitivity analysis to determine most influential parameters. Both first-order and total sensitivity indices profiles have been calculated throughout a time horizon of one year on main differential state variables, which include cyanobacteria, diatoms and chlorophyta concentration, as well as main nutrient concentration. Numerical results have shown a marked seasonality in sensitivity indices and have allowed ranking of parameters, which in turn have been estimated with observed data from the study case, Paso de las Piedras reservoir (Bahía Blanca, Argentina).

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Most water bodies in the world are becoming increasingly eutrophic due to anthropogenic activities. To address restoration, significant progress has been made, during the last decade, in the development of mechanistic water quality models (Hamilton and Schladow, 1997; Schladow and Hamilton, 1997; Omlin et al., 2001a; Zhang et al., 2004; Arhonditsis and Brett, 2005a,b). Once calibrated and validated, these models provide a useful tool to study ecosystem performance, to forecast biological communities response to changes in the driving conditions (climate change, alterations in nutrient loads) and to evaluate optimal restoration policies that can be applied to the water body (Anh et al., 2006; Saloranta and Andersen, 2007; Cossarini and Solidoro, 2008; Estrada et al., 2009b).

Mechanistic eutrophication models represent ecological processes through a set of complex nonlinear differential algebraic equations, with rate coefficients that require calibration to suit site-specific conditions (Hamilton and Schladow, 1997). Therefore, a dynamic parameter estimation problem is critical in the development of a eutrophication model. However, as in most cases it is not

possible to estimate all parameters due to their large number in relation to data availability, a ranking of the most influential parameters on model outputs must be made prior to parameter estimation.

In order to identify main parameters, that is, the parameters whose uncertainty gives larger variations in model outputs, sensitivity analysis is performed as an initial step in model development. Several sensitivity analysis methods have been applied in ecological modeling and they can be classified into two main categories: local and global. In local sensitivity analysis, the effect of infinitesimal variations one-at-a-time of the parameters on a particular model output is studied. These methods are based on a Taylor series linear expansion around the nominal trajectories and are model dependent (Omlin et al., 2001b; Pastres et al., 2003; Cariboni et al., 2007; Norton, 2008). Therefore, the local sensitivity approach can fail when the model under study is nonlinear, as it is the case in most ecological systems, with important interactions among input parameters (Pastres and Ciavatta, 2005; Cariboni et al., 2007).

Omlin et al. (2001b) applied a local sensitivity analysis described by Brun et al. (2002) to an ecological model developed for Lake Zürich. This approach involves a pre-selection of parameters that could be reasonably estimated from the available data and which are classified into three uncertainty categories based on previous knowledge (accurately, intermediate and poorly known parameters). The control parameter selection is based on diagnostic measures: the sensitivity and collinearity indices, which allow selection of parameter sets to use in calibration. This method has been applied to a few models

* Corresponding author at: PLAPIQUI, Planta Piloto de Ingeniería Química (UNS – CONICET), Camino La Carrindanga Km 7, Bahía Blanca, Argentina. Tel.: +54 2914861700; fax: +54 2914861600.

E-mail address: sdiaz@plapiqui.edu.ar (M.S. Diaz).

developed for aquatic ecosystems (e. g. Lindenschmidt, 2006; Anh et al., 2006; Mieleitner and Reichert, 2006).

On the other hand, global sensitivity analysis provides information on model outputs when the entire range of variation of parameters is explored simultaneously (Sobol', 1990, 2001; Sobol' and Levitan, 1999; Saltelli and Sobol', 1995; Homma and Saltelli, 1996; Saltelli et al., 1999, 2004; Saltelli and Tarantola, 2002). Although the disadvantage of these methods is its higher computational cost, the increasing computational power of computers allows the application of approaches based on sampling techniques such as Monte Carlo Simulations (MCS) or Latin Hypercube (LHS), even in complex models. They are very powerful, robust and flexible compared to the remaining methods (Manache and Melching, 2008). Furthermore, they provide more realistic results, since parameter interactions can be identified and are model-independent because the assumption of linearity or additivity of the model is not required (Saltelli et al., 2004). Håkanson (2000) has applied Monte Carlo Sampling to lake eutrophication models, as well as the characteristic coefficients of variation (CV) to study the uncertainty and sensibility in a set of lake variables. Manache and Melching (2008) have studied sensitivity for a model representing Denver River (Belgium) through Latin Hypercube Sampling technique and quantified the effects on model output with statistical measures based on regression and correlation analysis. The screening technique proposed by Morris (Morris, 1991; Campolongo et al., 2007; Nguyen and Kok, 2007) has been applied by Cossarini and Solidoro (2008) to a trophodynamic model of Gulf Trieste. Regarding the application of global sensitivity analysis to water quality models, Pastres et al. (1999) have studied the sensitivity of six factors in a model developed for the central part of Venice Lagoon. Chu et al. (2007) performed sensitivity analysis on a simple model of Black Sea consisting of three components (nutrients, phytoplankton and zooplankton), with twelve uncertain parameters. Another global sensitivity approach carried out on water quality models is Fourier Amplitude Sensitivity Test (FAST, Cukier et al., 1973, 1978; Koda et al., 1979a,b; McRae et al., 1982; Liepmann and Stephanopoulos, 1985; Sobol', 1990), which has been applied to MyLake model (Multi-year Lake Simulation, Saloranta and Andersen, 2007). In this method values for model parameters are sampled in a wave-like form, so that the amplitude of the particular wave is equal to the parameters predefined variation range.

In this work, we perform global sensitivity analysis to a first principle-based eutrophication model we have developed for Paso de las Piedras reservoir, which provides drinking water for two cities in Argentina (Bahia Blanca and Punta Alta) with a population of over 450,000. The model comprises twenty one differential equations and sixty algebraic ones, with sixty five parameters, from which we have considered twenty three as uncertain parameters. We have implemented a variance-based global sensitivity analysis, Sobol' method, to identify most influential model parameters on the water quality model as well as non-influential ones. Main parameters are estimated in another study, based on collected data from the reservoir (Estrada et al., 2009a) and the calibrated model is formulated within a dynamic optimization framework as an optimal control problem to determine bioremediation policies (Estrada et al., 2009b). The determination of dynamic Sobol'–Saltelli sensitivity indices (Saltelli et al., 1999; Sobol', 2001) has allowed ranking of the parameters according to their effect on main differential state variables variance.

2. Methods

2.1. First principle-based water quality model for Paso de las Piedras Reservoir

The case study is Paso de las Piedras Reservoir, which is located in Buenos Aires Province (Argentina) at 38° 22' S and 61° 12' W, within a protected area, Provincial

Park Paso de la Piedras. It is an artificial water body that was built by damming the Sauce Grande River to supply drinking water to more than 450,000 inhabitants of two cities in Argentina (Bahía Blanca and Punta Alta) and for industrial purposes at a petrochemical complex nearby. Main features of the reservoir are shown in Table 1. Another distinctive feature of this water body is that there is no stratification (Intartaglia and Sala, 1989), probably due to the fact that the reservoir is wind-exposed throughout the entire year. The reservoir has two tributaries, Sauce Grande River and El Divisorio Stream, which run through an important agricultural area in the country. The high discharge of plant nutrients (mainly phosphorus, nitrogen and silica) into the reservoir along with the high retention time (four years) has made this water body eutrophic. Therefore, there are recurrent cyanobacteria blooms that produce several problems to humans and the natural ecosystem (Intartaglia and Sala, 1989; Parodi et al., 2004).

To develop a first principle-based ecological water quality model of Paso de las Piedras Reservoir, we have formulated dynamic mass balances for the three main phytoplankton groups (Cyanobacteria, Diatoms and Chlorophyta), nutrients (nitrate, ammonium, organic nitrogen, phosphate and organic phosphate), dissolved oxygen and oxygen biochemical demand. We have included a correlation for zooplankton concentration profile, based on collected data from the reservoir, to take into account zooplankton grazing on phytoplankton. Main simplifying assumptions in the model are: horizontally averaged compositions, phosphorus as limiting nutrient and constant liquid density and reservoir transversal area. To consider horizontally averaged compositions, we analyzed collected data at four existing sampling stations. Two of them are coastal and the remaining ones are in central parts of the reservoir. The analysis of the data collected during a bloom period, gave up to 8, 9 and 12% average relative differences in nitrate, phosphate and total phytoplankton concentrations, respectively, as related to the first sampling station. These small differences can be explained with the fact that the reservoir is located in a windy area and it is wind-exposed throughout the entire year due to the low topology and support the simplifying assumption. A global mass balance takes into account not only inflows from tributaries and outflows to the potabilization plant and the downstream river, but inputs from rain (Q_{rain}) and outputs (Q_{evap}) due to evaporation (Eq. (1)). Composition gradients have been considered along the water column height, rendering a partial differential algebraic equations (PDAE) system. To transform the resulting PDAE into a set of ordinary differential equations, the column height is discretized into two layers, according to available observed data from the lake that can be used for parameter estimation Eqs. (2a) and (2b). Mass balances in each spatial layer include component inputs from tributaries (Q_{IN}), outputs for both potabilization and industrial purposes (Q_{OUT}) and the river itself, generation and consumption, and transference between layers, also accounting for lake volume variability (through upper layer height variability). Algebraic equations stand for the generation/consumption terms, as well as for seasonal climatological forcing functions. Forcing functions (temperature, solar radiation, rain, evaporation and river inflows and corresponding concentrations) are represented by sinusoidal functions, based on data from Paso de las Piedras Reservoir. Kinetic and physico-chemical parameters are included within the generation/consumption terms.

Total mass balance

$$\frac{dh_T}{dt} = \frac{1}{\rho A} \left[\sum_{k=1}^{N_{IN}} Q_{IN_k} - \sum_{m=1}^{N_{OUT}} Q_{OUT_m} + Q_{rain} - Q_{evap} \right], \quad (1)$$

where k stands for Sauce Grande River and El Divisorio Stream; m stands for outputs to potabilization and industrial purposes and the downstream Sauce Grande River.

Upper layer

$$\frac{dC_{Uj}}{dt} = \sum_{k=1}^{N_{IN}} \frac{Q_{IN_{Uk}}}{V_U} C_{IN_{Uk}} - \frac{Q_{OUT_U}}{V_U} C_{Uj} + r_{Uj} - \frac{k_d A}{\Delta h_U h_U} (C_{Uj} - C_{Lj}) - \frac{C_{Uj}}{h_U} \frac{dh_U}{dt} \quad (2a)$$

Lower layer

$$\frac{dC_{Lj}}{dt} = \frac{Q_{OUT_L}}{V_L} C_{Lj} + r_{Lj} + \frac{k_d A}{\Delta h_U h_U} (C_{Uj} - C_{Lj}) - \frac{C_{Lj}}{h_L} \frac{dh_L}{dt}, \quad (2b)$$

where U and L stand for the upper and lower layers, respectively; j , for cyanobacteria, diatoms, chlorophytes, ammonium, nitrate, phosphate, organic nitrogen, organic phosphorus, dissolved oxygen and carbonaceous biochemical demand of oxygen. Q_{OUT_U} and Q_{OUT_L} are the outputs to Sauce Grande River and potabilization plant and industrial purposes, respectively. Finally, h_U and V_U are lake upper layer

Table 1
Paso de las Piedras Reservoir data (Schefer, 2004).

Area of drainage basin (km ²)	1620
Perimeter of coastline (km)	60
Surface (km ²)	36
Mean depth (m)	8.2 m
Maximum depth (m)	28 m
Maximum volume (hm ³)	328 Hm ³
Retention time (years)	4

Table 2Rate equations for Paso de las Piedras eutrophication model, Eqs. (2a) and (2b); i = upper (U) and lower layer (L); j = cyanobacteria, diatoms and chlorophytes.

Phytoplankton growth	$R_{ij} = R_{ij, \text{growth}} - R_{ij, \text{resp}} - R_{ij, \text{death}} - R_{ij, \text{settling}} - R_{ij, \text{graz}}$
Phytoplankton growth rate	$R_{ij, \text{growth}} = k_{i, \text{growth}} f(T)_{ij} f(I)_{ij} f(N)_{ij} C_{ij}$
Temperature limitation	$f(T)_{ij} = -\frac{(T_j - T_{\text{opt}_i})^2}{T_{\text{opt}_i}^2} + 1$
Nutrient limitation	$f(N)_{ij} = \begin{cases} \frac{C_{i, \text{PO}_4}}{C_{i, \text{PO}_4} + K_{P_i}} \cdot \frac{C_{\text{Si}_i}}{C_{\text{Si}_i} + K_{S_i}} & i = D \\ \frac{C_{\text{PO}_4}}{C_{\text{PO}_4} + K_{P_i}} & i = C, G \end{cases}$
Light limitation	$f(I)_{ij} = \frac{I_{oi}}{I_{\text{opt}_j}} \exp\left(1 - \frac{I_{oi}}{I_{\text{opt}_j}}\right)$
Respiration	$R_{ij, \text{resp}} = k_{j, \text{resp}} \theta_r^{(T-20)} C_{ij}$
Death	$R_{ij, \text{death}} = k_{j, \text{death}} \theta_m^{(T-20)} C_{ij}$
Settling	$R_{ij, \text{settling}} = k_{j, \text{settling}} \frac{C_{ij}}{h_i}$
Grazing	$R_{ij, \text{graz}} = k_{j, \text{graz}} \frac{C_{ij}}{C_{ij} + K_{\text{graz}}} C_{\text{Zoo}_i}$
Rate of change of nitrogen cycle compounds concentration	
r_{U, NH_4}	$-R_{U, \text{NH}_4, \text{uptake}} - R_{iU, \text{nit}} + R_{U, \text{NH}_4, \text{death}} + R_{U, \text{NH}_4, \text{miner}}$
r_{L, NH_4}	$-R_{L, \text{NH}_4, \text{uptake}} - R_{Li, \text{nit}} + R_{L, \text{NH}_4, \text{death}} + R_{L, \text{miner}} + R_{L, \text{NH}_4, \text{sediment}}$
r_{i, NO_3}	$-R_{i, \text{NO}_3, \text{uptake}} + R_{i, \text{NO}_3, \text{nit}} - R_{i, \text{NO}_3, \text{denit}}$
$r_{i, \text{ON}}$	$R_{i, \text{ON}, \text{death}} - R_{i, \text{miner}} - R_{i, \text{ON}, \text{settling}}$
Phytoplankton NH_4 uptake	$R_{i, \text{NH}_4, \text{uptake}} = \sum_j (\alpha_{\text{nc}} R_{j, \text{growth}} P_{\text{NH}_4})$
NH_4 preference factor	$P_{\text{NH}_4} = 1 - \exp(-\psi_j C_{i, \text{NH}_4})$
NH_4 release by phytoplankton death	$R_{i, \text{NH}_4, \text{death}} = \sum_j (C_{ij} k_{j, \text{death}} \alpha_{\text{nc}} (1 - f_{\text{ON}}))$
Nitrification rate	$R_{i, \text{NH}_4, \text{nit}} = k_{\text{nit}} \theta_{\text{nit}}^{(T-20)} \frac{C_{i, \text{DO}}}{K_{\text{nit}} + C_{i, \text{DO}}} C_{i, \text{NH}_4}$
ON mineralization rate	$R_{i, \text{miner}} = k_{\text{mn}} \theta_{\text{mn}}^{(T-20)} \frac{\sum_j C_{ij}}{K_{\text{mpc}} + \sum_j C_{ij}} C_{i, \text{ON}}$
NH_4 release from sediments	$R_{L, \text{NH}_4, \text{sediment}} = S_{\text{N}} \left(1 - \frac{C_{L, \text{DO}}}{K_{\text{DOS}} + C_{L, \text{DO}}}\right) A$
Phytoplankton NO_3 uptake	$R_{i, \text{NO}_3, \text{uptake}} = \sum_j (\alpha_{\text{nc}} R_{j, \text{growth}} (1 - P_{\text{NH}_4}))$
Denitrification rate	$R_{i, \text{NO}_3, \text{denit}} = k_{\text{denit}} \theta_{\text{denit}}^{(T-20)} \frac{K_{\text{denit}}}{K_{\text{denit}} + C_{i, \text{DO}}} C_{i, \text{NO}_3}$
ON release by phytoplankton death	$R_{i, \text{ON}, \text{death}} = \sum_j (C_{ij} k_{j, \text{death}} \alpha_{\text{nc}} f_{\text{ON}})$
ON settling rate	$R_{i, \text{ON}, \text{settling}} = \frac{k_{\text{ON}, \text{settling}} (1 - f_{\text{DON}})}{D_i} C_{i, \text{ON}}$
Rate of change of phosphorus cycle compounds concentration	
r_{U, PO_4}	$-R_{U, \text{PO}_4, \text{uptake}} + R_{U, \text{PO}_4, \text{death}} + R_{U, \text{miner}}$
r_{L, PO_4}	$-R_{L, \text{PO}_4, \text{uptake}} + R_{L, \text{PO}_4, \text{death}} + R_{L, \text{miner}} + R_{L, \text{PO}_4, \text{sediment}}$
$r_{i, \text{OP}}$	$R_{i, \text{OP}, \text{death}} - R_{ij, \text{miner}} - R_{i, \text{OP}, \text{settling}}$
Phytoplankton PO_4 uptake	$R_{i, \text{PO}_4, \text{uptake}} = \sum_j (\alpha_{\text{pc}} R_{j, \text{growth}})$
PO_4 release by phytoplankton death	$R_{i, \text{PO}_4, \text{death}} = \sum_j (C_{ij} k_{j, \text{death}} \alpha_{\text{pc}} (1 - f_{\text{OP}}))$
PO_4 release from sediments	$R_{L, \text{PO}_4, \text{sediment}} = S_{\text{P}} \left(1 - \frac{C_{L, \text{DO}}}{K_{\text{DOS}} + C_{L, \text{DO}}}\right) A$
OP release by phytoplankton death	$R_{i, \text{OP}, \text{death}} = \sum_j (C_{ij} k_{j, \text{death}} \alpha_{\text{pc}} f_{\text{OP}})$
OP mineralization rate	$R_{i, \text{miner}} = k_{\text{mp}} \theta_{\text{mp}}^{(T-20)} \frac{\sum_j C_{ij}}{K_{\text{mpc}} + \sum_j C_{ij}} C_{i, \text{OP}}$
OP settling rate	$R_{i, \text{OP}, \text{settling}} = \frac{k_{\text{OP}, \text{settling}} (1 - f_{\text{DOP}})}{D_i} C_{i, \text{OP}}$
Rate of change of dissolved oxygen concentration	
$r_{U, \text{DO}}$	$R_{U, \text{DO}, \text{reair}} - R_{U, \text{DO}, \text{nitr}} + R_{U, \text{DO}, \text{resp/phot}} - R_{U, \text{DO}, \text{bod}}$
$r_{L, \text{DO}}$	$-R_{L, \text{DO}, \text{nitr}} - R_{L, \text{DO}, \text{resp/phot}} - R_{L, \text{DO}, \text{bod}} + R_{L, \text{DO}, \text{sediment}}$
Re-aeration rate	$R_{U, \text{DO}, \text{reair}} = \frac{k_a \theta_a^{(T-20)}}{D_U} (C^* - C_S)$
Saturated dissolved oxygen	$C^* = 16.5 - \frac{8}{22} T$
DO consumed in nitrification process	$R_{i, \text{DO}, \text{nitr}} = k_{\text{nit}} \theta_{\text{nit}}^{(T-20)} \frac{C_{i, \text{DO}}}{K_{\text{nit}} + C_{i, \text{DO}}} C_{i, \text{NH}_4} \alpha_{\text{on}}$
Net production of oxygen by respiration/photosynthesis process	$R_{i, \text{DO}, \text{resp/phot}} = (\sum_j R_{ij, \text{growth}} - R_{ij, \text{resp}}) \alpha_{\text{oc}}$
DO for organic matter oxidation	$R_{i, \text{DO}, \text{bod}} = k_{\text{bod}} \theta_{\text{bod}}^{(T-20)} \frac{C_{i, \text{DO}}}{K_{\text{bod}} + C_{i, \text{DO}}} C_{\text{BDO}_i}$
Sediment oxygen demand	$R_{L, \text{DO}, \text{sediment}} = \frac{k_{\text{sod}} \theta_{\text{sod}}^{(T-20)}}{D_L} \frac{C_{L, \text{DO}}}{K_{\text{sod}} + C_{L, \text{DO}}}$
Zooplankton concentration correlation	$C_{\text{Zoo}_i} = 6.79 \times 10^{-4} t^6 - 2.65 \times 10^{-2} t^5 + 3.93 \times 10^{-1} t^4 - 2.74 t^3 + 8.80 t^2 - 10.26 t + 1.12$

height and volume, respectively, while A is the reservoir transversal area and k_d is the vertical eddy diffusivity rate. Table 2 shows main rate equations, r_{Uj} and r_{Lj} , in Eqs. (2a) and (2b) (Estrada et al., 2009a). Table 3 summarizes the entire set of model parameters and their nominal value.

2.2. Global sensitivity analysis on eutrophication models

Sensitivity analysis approaches can be broadly classified into local and global ones. Local techniques compute sensitivity indices as the first partial derivative of model output variables with respect to the parameter of interest. They are based on a Taylor series expansion around the parameter nominal value. However, the assumption of linearity is valid only within a narrow range of variation of the parameter. Therefore, results from local sensitivity analysis may not be representative when the entire range of variation of parameters must be considered and when dealing with nonlinear models, such as complex eutrophication mechanistic ones. On the other hand, global sensitivity analysis is based on exploring the total range of variation of model parameters, sampling from the distribution function associated to each input parameter and on performing repeated model simulations. These methods have a higher computational cost than local techniques, but they provide more realistic results, since parameter interactions can be identified (Saltelli et al., 2004; Pastres and Ciavatta, 2005). Furthermore, global methods do not require the assumption of model linearity and/or additivity. State of the art techniques for global sensitivity analysis include Morris method (Morris, 1991), Fourier Amplitude Sensitivity Test (FAST) (Cukier et al., 1973) and Sobol' method.

Sobol' method is based on the same decomposition of variance as FAST, but through the application of Monte Carlo methods in place of spectral analysis (Sobol', 1990; Saltelli and Sobol', 1995; Sobol', 2001; Saltelli and Tarantola, 2002). The basic idea is that given a function $y = f(x, t)$, where y is a state differential or algebraic variable (e.g. phytoplankton or nutrient concentration in the eutrophication model) and x is a vector of k model input parameters (e.g. any model parameter from Table 2) and t is time, this function may be decomposed into terms of increasing dimensionality (Sobol', 1990). For the sake of clarity, in the following discussion we drop the index t , assuming f and its expected value and variance are calculated at each time instant:

$$f = f_0 + \sum_i f_i(x_i) + \sum_{i>j} f_{ij}(x_i, x_j) + \dots + f_{12\dots k}(x_1, x_2, \dots, x_k) \tag{3}$$

Eq. (3) is an ANOVA representation of the function $y = f(x)$ if

$$\int f_{i_1\dots i_s}(x_{i_1}\dots x_{i_s}) dx_{i_1}\dots dx_{i_s} = 0 \tag{4}$$

From (4), the summands in Eq. (3) are orthogonal and can be expressed as integrals of $f(x)$. By squaring Eq. (3) and integrating:

$$\int f^2(x) dx - f_0^2 = \sum_{s=1}^k \sum_{i_1 < \dots < i_s} f_{i_1\dots i_s}^2 dx_{i_1}\dots dx_{i_s} \tag{5}$$

where

$$V(y) = \int f^2(x) dx - f_0^2, \quad V_{i_1\dots i_s} = \int f_{i_1\dots i_s}^2 dx_{i_1}\dots dx_{i_s} \tag{6}$$

$V(y)$ and correspond to the unconditional and conditional variance of the state variable, respectively. The unconditional variance can be written within the variance-decomposition scheme (Sobol', 1990), as:

$$V(y) = \sum_i V_i + \sum_{i>j} V_{ij} + \dots + V_{12\dots k} \tag{7}$$

Eqs. (3) and (6) are unique if input parameters are orthogonal and the summands in Eq. (3) are square integrable in the domain of existence.

Despite that input factors are orthogonal or not, the unconditional variance can be decomposed into the sum of the variance of a conditional expected value and the expected value of a conditional variance:

$$V(y) = V(E(y|x_i)) + E(V(y|x_i)) \tag{8}$$

$$V(y) = V(E(y|x_{-i})) + E(V(y|x_{-i})) \tag{9}$$

where V and E correspond to variance and expected value operators, respectively. In Eq. (8), $V(E(y|x_i)) = V_i$ computes the variance (overall possible realizations of parameter x_i) of the conditional expected value of the state variable y under all parameters variation, except x_i . It represents the expected reduction in the state variable variance could be obtained if x_i could be known or fixed. It is the first-order effect associated to parameter x_i . The second term, $E(V(y|x_i))$, is the expected value (overall realizations of parameter x_i) of the conditional variance of the state variable y under all parameters variation, except x_i . It represents the average state variable variance that would be left if x_i could be known or could be fixed.

The same can be stated for Eq. (9), by replacing x_i for "all parameters except x_i " (x_{-i}). In this way, the term $E(V(y|x_{-i})) = V_i^{TOT}$ computes the average state variable

variance that would be left if all parameters except x_i could be known or could be fixed; i.e., it takes into account all the terms in Eq. (3) that include x_i , for the case of orthogonal input factors.

If Eqs. (8) and (9) are divided by the unconditional variance, the following expressions are obtained:

$$1 = \frac{V(E(y|x_i))}{V(y)} + \frac{E(V(y|x_i))}{V(y)} \tag{10}$$

$$1 = \frac{V(E(y|x_{-i}))}{V(y)} + \frac{E(V(y|x_{-i}))}{V(y)} \tag{11}$$

The first-order sensitivity index, S_i and the total sensitivity index S_i^{TOT} are defined as:

$$S_i = \frac{V(E(y|x_i))}{V(y)} = \frac{V_i}{V(y)} \tag{12}$$

$$S_i^{TOT} = \frac{E(V(y|x_{-i}))}{V(y)} = \frac{V_i^{TOT}}{V(y)} \tag{13}$$

Both S_i and S_i^{TOT} are good measures for the effect of variation in parameters on model outputs. S_i gives the reduction on the unconditional variance of the state variable that can be obtained if x_i is fixed at its true value. On the other hand, S_i^{TOT} takes into account the interactions among parameters, so it gives information on the non-additive part of the model.

Usually, $\sum_{i=1}^k S_i < 1$ and $S_i < S_i^{TOT}$. However for a purely additive model and orthogonal inputs, $\sum_{i=1}^k S_i = 1$, which can be observed if Eq. (7) is divided by $V(y)$ and the interaction terms are canceled.

An additional index, S_i^{int} takes into account the effects of interactions among model parameters and it can be calculated as:

$$S_i^{int} = S_i^{TOT} - S_i \tag{14}$$

In summary, when S_i is high, x_i is an influential parameter; when S_i and S_i^{TOT} are very different, there are important interactions between x_i and other parameters (S_i^{int} is important); when S_i and S_i^{TOT} are nearly equal (S_i^{int} nearly zero), there is little or no interaction between x_i and other parameters; when both S_i and S_i^{TOT} are small, x_i is not an influential parameter (neither alone nor in interactions with other parameters).

Sobol' (2001) has proposed a methodology to compute sensitivity indices, based on Monte Carlo simulations, with a minimum amount of function evaluations. Main steps for the dynamic case are as follows:

1. Generation of two different random sets of model parameters: $\xi = (\eta, \zeta)$ and $\xi' = (\eta', \zeta')$ at each time instant. Each matrix has dimension $N \times k$, where N is the sample size for the Monte Carlo method and k is the number of parameters; η is a vector of dimension $N \times 1$, which contains N random values for parameter x_i whose sensitivity indices are to be calculated, and ζ is a submatrix of dimension $N \times (k - 1)$ and contains random values for the $k - 1$ remaining input parameters.
2. Generation of two new matrices by combination of ξ and ξ' , which are required for the computation of the time dependent expected values for state variables, compositions, ($c_0(t)$ stands for $E(c(t)|x_i)$) and unconditional variances ($V(t)$), at each time instant t , as well as the conditional variances ($V(t)$; stands for $V(E(c(t)|x_i))$ and $V(t)_{-i}$ stands for $V(E(c(t)|x_{-i}))$):

$$c_0(t) = \frac{1}{N} \sum_{i=1}^N c(t, \xi_i) \tag{15}$$

$$V(t) = \frac{1}{N} \sum_{i=1}^N c^2(t, \xi_i) - c_0^2(t) \tag{16}$$

$$V(t)_i = \frac{1}{N} \sum_{i=1}^N c(t, \xi_i) c(t, \eta_i, \zeta'_i) - c(t, \xi_i) c(t, \xi'_i) \tag{17}$$

$$V(t)_{-i} = \frac{1}{N} \sum_{i=1}^N c(t, \xi_i) c(t, \eta'_i, \zeta_i) - c_0^2(t) \tag{18}$$

3. Sensitivity indices profiles are calculated by the corresponding definitions given above throughout the time horizon.

2.3. Numerical experiments

In the present study we have applied Global Sensitivity Analysis (GSA), as described in the previous section, to the ecological water quality model of Paso de las

Piedras Reservoir, Eqs. (1) and (2) and equations from Table 2 (Estrada et al., 2009a). Global sensitivity analysis on a large-scale differential algebraic system of equations has required special computational effort. As a first step, we have associated normal probability distributions to each parameter as these distributions have been frequently associated to parameters from an ecological water model in the literature (Baklouti et al., 2006; Pastres and Ciavatta, 2005; Manache and Melching, 2008). Regarding probability distribution parameters, we have assigned 10% standard deviation around the nominal value for each parameter throughout the entire time horizon to run the global sensitivity analysis. We have also explored changing standard deviations according to variability of parameter values in the literature, with standard deviations shown in Table 4. However, numerical results have shown low sensitivity to these changes and do not change the parameter ranking. As a second step, Monte Carlo simulations have been performed for the three matrices of random parameters, as described in Section 2.2. The required number of scenarios N is function of model complexity (Baklouti et al., 2006). The number of scenarios has been estimated by performing stochastic simulations for increasing number of scenarios and comparing mean value profiles for main differential variables. As it can be seen in Figs. 1 and 2 for cyanobacteria and phosphate concentrations, both mean concentration profiles remain unchanged for 1700–2000 scenarios, suggesting that considering 1700 scenarios is appropriate for this model. The set of twenty three parameters (Table 4) has been previously selected by performing stochastic simulations when parameters vary one-at-a time over their range of variation, and calculating the unconditional variance of main state variables, so parameters that were feasible for tuning under biological considerations and produce values for variances greater than 0.5 were selected. We have studied parameter influence on the key state variables for the water quality model of Paso de las Piedras Reservoir, i.e., cyanobacteria, diatoms, chlorophytes, nitrate and phosphate concentrations. Stochastic simulations have been carried out in gPROMS (Process Systems Enterprise, 2009) and numerical results have been exported to calculate sensitivity indices within a Fortran 90 environment, as they could not be carried out in gPROMS due to model size. Finally, we have calculated Sobol' first and total order sensitivity indices profiles for the twenty three selected parameters, as Saltelli et al. (1999) and Sobol' (2001), Eqs. (15)–(18) and (12)–(14), on a daily basis, within a time horizon of one year.

3. Results

Global Sensitivity Analysis numerical results are presented in Figs. 3–12 and Tables 5–10. Figs. 3, 5, 7, 9 and 11 show the temporal variation for first-order sensitivity indices (S_i) for cyanobacteria, diatoms, chlorophytes, nitrate and phosphate concentrations, respectively. An important feature of the dynamic sensitivity analysis is that the influential parameter set changes with time, which indicates changes in the dynamic behavior of the model throughout the time horizon. As it can be seen in these figures, the sum of S_i is seldom equal to 1. Because the summation of the first-order sensitivity indices is a measure of model additivity (Saltelli et al., 2004), the cumulative area plots of S_i show that there are important interactions among parameters for the output of the five analyzed state variables. We have also calculated S_i^{int} in order to quantify the contribution of parameter interaction. As indicated in Section 2.2, this index is obtained as $(S_i^{\text{TOT}} - S_i)$ and it is a measure of how much a parameter is involved in interactions with any other input parameter (Figs. 4, 6, 8, 10 and 12). Tables 5–9 show S_i and S_i^{int} values at four different time periods of the year corresponding to phytoplankton peaks for parameters included in the present sensitivity analysis.

Cyanobacteria concentration is the state variable that has larger values for $\sum S_{i=1, \dots, 23}$ throughout the entire time horizon, ranging from 0.19 to 1.00 (Fig. 3 and Table 5). In summer, most of the parameters have influence on cyanobacteria concentration profile, but the main ones are θ_{mp} (temperature factor for organic phosphorus mineralization rate), θ_r (temperature factor for phytoplankton respiration rate), θ_m (temperature factor for phytoplankton mortality rate), and K_1 (background light attenuation), as indicated by their larger S_i values. Furthermore, the summation over these four parameters explains around 44% of cyanobacteria concentration variance. The remaining parameters have less influence during this season. In fall and winter, main parameters are θ_{mp} , θ_r , θ_m and K_1 , explaining a mean variance of 56%, ranging between 20 and 86%. In spring, f_{ON} (fraction of dead

and respired phytoplankton recycled to organic nitrogen pool), α_{nc} (phytoplankton nitrogen to carbon ratio) and I_{optC} (optimal light intensity for cyanobacteria) also become important, with a mean variance of 45%. With respect to interactions among parameters, Fig. 4 and Table 5 show that $k_{\text{C,growth}}$ (maximum growth rate for cyanobacteria) is the most influential parameter over the entire year, which indicates that even though $k_{\text{C,growth}}$ is not important in first effect terms, it does affect cyanobacteria concentration through interactions with other input parameters. In summer, $k_{\text{C,growth}}$ has the largest S_i^{int} value and except for f_{ON} , k_{mn} , I_{optD} (optimal light intensity for diatoms), θ_{mp} , θ_r , θ_m , the remaining parameters are only influential through interactions with other parameters. During the fall, $k_{\text{C,growth}}$ is the only important factor, while in winter and spring, θ_r , $k_{\text{D,death}}$, I_{optC} , K_1 also become significant.

Regarding S_i profiles for diatom concentration, a marked seasonality can be observed, with an average summation of 0.29, ranging from 0 to 1. As with cyanobacteria concentration, several parameters influence diatoms concentration variance in summer and the $\sum S_{i=1, \dots, 23}$ is equal to one at three times throughout the time horizon (Fig. 5 and Table 6). However, in fall first effect indices are not relevant while parameter interactions in this season are very important (Fig. 6). The most important factors in this period are $k_{\text{G,growth}}$ (maximum growth rate for chlorophytes), θ_{mn} and K_1 . In winter, first effect terms become more important reaching a $\sum S_{i=1, \dots, 23}$ value of 0.71 with the following main parameters: θ_{mp} , θ_r and $k_{\text{G,growth}}$. Whereas in early spring these factors remain important, toward the middle of the season, the set of influential factors changes and important parameters are θ_m , $k_{\text{C,growth}}$ and K_1 . Finally, we can conclude that over the entire year the set of most influential factors related to parameter interactions includes $k_{\text{G,growth}}$, θ_{mp} , θ_r and θ_m (Fig. 6 and Table 6).

Average, minimum and maximum values of $\sum S_{i=1, \dots, 23}$ for chlorophytes are 0.23, 0 and 1, respectively. As with the other phytoplankton groups (cyanobacteria and diatoms), in summer $\sum S_{i=1, \dots, 23}$ has larger values and several parameters are important (Fig. 7 and Table 7). Parameters that have the most significant effects on chlorophytes concentration are θ_{mp} and $k_{\text{G,growth}}$ in fall, and $k_{\text{D,growth}}$ and $k_{\text{C,growth}}$ in winter. They account for 1–20% and 0.3–11% of the chlorophytes concentration variance, respectively. By mid spring, first-order effects become more important explaining up to 49% of the variance and main parameters are K_1 , θ_m and $k_{\text{C,growth}}$. S_i^{int} for chlorophytes show that there is a regular set of parameters that influence this state variable through interactions over the entire time horizon, with major influence in winter and early spring. Influential parameters are θ_{mp} , θ_r , θ_m , K_{nit} and $k_{\text{G,growth}}$. In addition, $k_{\text{C,growth}}$ becomes important in fall, k_{mn} , I_{optG} , I_{optD} , $k_{\text{G,settling}}$ and $k_{\text{G,death}}$ in winter and K_1 in spring (Fig. 8 and Table 7).

In summary, the most important parameters for the three phytoplankton groups are: temperature adjustment for phosphorus mineralization rate, phytoplankton death and respiration and the background light attenuation coefficient (θ_{mp} , θ_r , θ_m and K_1). It is important to note that cyanobacteria concentration is not affected by its own maximum growth rate or any other (diatoms or chlorophytes maximum growth rate) in first-order effects, while diatoms and chlorophytes concentration are greatly influenced by their maximum growth rate. Also, diatoms and chlorophytes parameters are those that are most involved in interactions.

Regarding nitrate concentration, the parameters with most significant effects in summer are $k_{\text{G,growth}}$, θ_{mn} , α_{nc} , I_{optG} and $k_{\text{D,settling}}$, and $k_{\text{C,growth}}$ (Fig. 9). These parameters explain between 14 and 100% of the entire nitrate concentration variability. K_{nit} (half-saturation constant for oxygen limitation of nitrification) is the most important parameter in fall and it is still influential until the middle of spring, when θ_{mn} , θ_r and K_1 are also important. The set

Table 3
Model parameters for Paso de las Piedras eutrophication model and their nominal value.

Parameter	Description	Value
$k_{C,growth}$ (day^{-1})	Maximum growth rate for cyanobacteria	0.195
$k_{D,growth}$ (day^{-1})	Maximum growth rate for diatoms	0.653
$k_{G,growth}$ (day^{-1})	Maximum growth rate for chlorophytes	0.785
$k_{C,death}$ (day^{-1})	Mortality rate for cyanobacteria	0.019
$k_{D,death}$ (day^{-1})	Mortality rate for diatoms	0.100
$k_{G,death}$ (day^{-1})	Mortality rate for chlorophytes	0.100
$k_{C,resp}$ (day^{-1})	Respiration rate for cyanobacteria	0.047
$k_{D,resp}$ (day^{-1})	Respiration rate for diatoms	0.43
$k_{G,resp}$ (day^{-1})	Respiration rate for chlorophytes	0.42
$k_{C,settling}$ (m/day)	Settling velocity for cyanobacteria	0.15
$k_{D,settling}$ (m/day)	Settling velocity for diatoms	0.20
$k_{G,settling}$ (m/day)	Settling velocity for chlorophytes	0.15
T_{optC} ($^{\circ}C$)	Optimal temperature for cyanobacteria	30
T_{optD} ($^{\circ}C$)	Optimal temperature for diatoms	19
T_{optG} ($^{\circ}C$)	Optimal temperature for chlorophytes	20
I_{optC} (ly/day)	Optimal light intensity for cyanobacteria	99
I_{optD} (ly/day)	Optimal light intensity for diatoms	25
I_{optG} (ly/day)	Optimal light intensity for chlorophytes	89.8
K_{PC} (mgP/l)	Half-saturation constant for P uptake by cyanobacteria	0.0002
K_{PD} (mgP/l)	Half-saturation constant for P uptake by diatoms	0.05
K_{PG} (mgP/l)	Half-saturation constant for P uptake by chlorophytes	0.0009
K_S (mgSi/l)	Half-saturation constant for Si uptake	0.0053
k_{nit} (day^{-1})	Nitrification rate	0.09
k_{denit} (day^{-1})	Denitrification rate	0.001
k_{mn} (day^{-1})	Organic nitrogen mineralization rate	0.032
k_{mp} (day^{-1})	Organic phosphorus mineralization rate	0.02
k_{bod} (day^{-1})	CBOD deoxygenation rate	0.5
k_{sod} (day^{-1})	Sediment oxygen demand rate	0.65
S_N (gN/m ² day)	Release rate of ammonium from the sediment	0.4
S_P (gP/m ² day)	Release rate of phosphate from the sediment	0.013
K_{nit} (mg/day)	Half-saturation const. for oxygen limitation of nitrification	0.115
K_{denit} (mg/day)	Half-saturation const. for oxygen limitation of denitrification	0.2
K_{bod} (mg/day)	Half-saturation const. for oxygen limitation of CBOD oxidation	0.5
K_{sod} (mg/day)	Half-saturation const. for sediment oxygen demand	0.4
K_{mpc} (mg/day)	Half-saturation const. for phytoplankton limitation	1.00
K_{DOS} (mg/day)	Half-saturation const. for nutrient sediment fluxes	0.4
$k_{ON,settling}$ (m/day)	Settling velocity for organic nitrogen	0.03
$k_{OP,settling}$ (m/day)	Settling velocity for organic phosphorus	0.03
$K_{bod,settling}$ (m/day)	Settling velocity for organic CBOD	0.03
θ_{ni}	Temperature adjustment for nitrification rate	1.080
θ_{denit}	Temperature adjustment for denitrification rate	1.080
θ_N	Temperature adjustment for release of NH ₄ sediment rate	1.080
θ_P	Temperature adjustment for release of PO ₄ sediment rate	1.080
θ_{bod}	Temperature adjustment for CBOD deoxygenation rate	1.050
θ_r	Temperature adjustment for phytoplankton respiration rate	1.04
θ_m	Temperature adjustment for phytoplankton mortality rate	1.02
θ_{mn}	Temperature adjustment for ON mineralization rate	1.08
θ_{mp}	Temperature adjustment for OP mineralization rate	1.10
θ_{sod}	Temperature adjustment for oxygen sediment demand rate	1.08
f_{ON}	Fraction of dead and respired phytoplankton recycled to ON pool	0.5
f_{OP}	Fraction of dead and recycled phytoplankton recycled to OP pool	0.5
f_{DON}	Fraction of dissolved organic nitrogen	1.00
f_{DOP}	Fraction of dissolved organic phosphorus	1.00
f_{DBOD}	Fraction of dissolved CBOD	1.00
α_{oc} (mgO/mgC)	Oxygen to carbon ratio	2.67
$cchl$ (mgC/mgChl)	Phytoplankton carbon to chlorophyll ratio	50
K_2 (m ² /mg)	Light attenuation coefficient for chlorophyll	0.002
$\Psi_{C,D,G}$ (mgN/m ³) ⁻¹	Strength of the ammonium preference	0.009
k_a (day^{-1})	Re-aeration rate	0.38
α_{on} (mgO/mgN)	Stoichiometric oxygen to nitrogen ratio for nitrification	0.216
α_{nc} (mgN/mgC)	Phytoplankton nitrogen to carbon ratio	0.125
α_{pc} (mgP/mgC)	Phytoplankton phosphorus to carbon ratio	0.01
K_1 (m ⁻¹)	Background light attenuation	1.97
A (km ²)	Transversal area of the lake	36
K_d (m ² day ⁻¹)	Vertical eddy diffusivity rate	0.543

of main parameters involved in interactions with others parameters for nitrate concentration include θ_r , θ_m , θ_{mp} , $k_{G,growth}$ and $k_{C,growth}$. For winter and spring, θ_{mn} is also influential (Fig. 10 and Table 8).

θ_r , θ_m , θ_{mp} , $k_{G,growth}$ and $k_{C,growth}$ have the most significant effects on phosphate concentration (Fig. 11 and Table 9) and they account

for about 32% of the overall observed variability. Until mid summer, first-order effects explain the total variance of phosphate concentration. Throughout fall, $k_{G,growth}$ is the only parameter with first-term index accounting for about 24% of the output variance, while $k_{G,growth}$, θ_m and θ_{mp} account for 35% in winter. In summer, the average variability related to $\sum S_{i=1,\dots,23}$ is 26% and the set of

Table 4

List of parameters used for GSA and equations in which they are involved.

Parameter	St. dev.	Variable	Parameter	St. dev.	Variable
α_{nc}	0.10	NO_{3j}, NH_{4j}, ON_j	$k_{G,death}$	0.10	$G_j, ON_j, NH_{4j}, PO_{4j}, OP_j$
α_{pc}	0.10	PO_{4j}, OP_j	θ_m	0.05	$C_j, D_j, G_j, ON_j, NH_{4j}, PO_{4j}, OP_j, BOD_j$
f_{ON}	0.10	NO_{3j}, NH_{4j}, ON_j	θ_r	0.05	C_j, D_j, G_j, DO_j
K_1	0.10	$C_j, D_j, G_j, NO_{3j}, NH_{4j}, PO_{4j}, DO_j$	θ_{mn}	0.05	NH_{4j}, ON_j
K_{mn}	0.10	NH_{4j}, ON_j	θ_{mp}	0.05	PO_{4j}, OP_j
K_{mp}	0.10	PO_{4j}, OP_j	$k_{C,growth}$	0.30	$C_j, NO_{3j}, NH_{4j}, PO_{4j}, DO_j$
k_{nit}	0.10	NO_{3j}, NH_{4j}, DO_j	$k_{D,growth}$	0.30	$D_j, NO_{3j}, NH_{4j}, PO_{4j}, DO_j$
K_{nit}	0.10	NO_{3j}, NH_{4j}	$k_{G,growth}$	0.30	$G_j, NO_{3j}, NH_{4j}, PO_{4j}, DO_j$
I_{optC}	0.10	$C_j, NO_{3j}, NH_{4j}, PO_{4j}, DO_j$	$k_{C,settling}$	0.10	C_j
I_{optD}	0.10	$D_j, NO_{3j}, NH_{4j}, PO_{4j}, DO_j$	$k_{D,settling}$	0.10	D_j
I_{optG}	0.10	$G_j, NO_{3j}, NH_{4j}, PO_{4j}, DO_j$	$k_{G,settling}$	0.10	G_j
$k_{D,death}$	0.10	$D_j, ON_j, NH_{4j}, PO_{4j}, OP_j$			

influential parameters includes $k_{G,growth}$, $k_{C,growth}$ and θ_m . The maximum growth rate of the three phytoplankton groups ($k_{C,growth}$, $k_{G,growth}$ and $k_{D,growth}$), θ_{mp} , θ_r and I_{optD} are the main parameters involved in interactions for phosphate concentration (Fig. 12 and Table 9).

Finally, Table 10 provides a summary of the contribution of the selected set of main parameters to account for cyanobacteria and phosphate concentration variability. It can be seen that the first six parameters (θ_r , θ_m , θ_{mp} , $k_{C,growth}$, $k_{G,growth}$ and K_1) account for the entire variability of phosphate concentration along the entire year, while they account for 70% of cyanobacteria concentration during the summer bloom and almost 100% of its variability during the rest of the year (as previously discussed, $k_{C,growth}$ does not account as first-order index, but through interactions).

4. Discussion

The global variance-based sensitivity analysis applied in this work aims at establishing the relative importance of the parameters involved in the seasonal dynamic of a eutrophication model aiming at two major issues:

- 1) Parameters fixing setting, to identify the parameters that can be fixed at any given value in their domains without significantly reducing the state variables variance (Homma and

Saltelli, 1996). In this case, we study total effect sensitivity indices, and

- 2) Parameters prioritization setting, to determine the most important factors which, if fixed to their true value, could lead to the greatest reduction in state variables variance (Saltelli and Tarantola, 2002). First effect sensitivity indices provide a reliable measure in this case.

Even though we are presenting a case study, the methodology is quite general and we show how main parameters can be determined in a complex mechanistic model for water quality. From factors fixing setting analysis, we have removed six of twenty three parameters. One of them is a stoichiometric parameter (α_{pc}), three are related to nutrient kinetics (k_{mp} , k_{mn} , k_{nit}) and the remaining two are settling rates for cyanobacteria and diatoms ($k_{C,settling}$ and $k_{D,settling}$), respectively. From the remaining parameters, k_{mn} , I_{optD} and $k_{C,settling}$ are engaged only in interactions with other parameters, while the others are mainly involved in first effect terms.

Based on first effect sensitivity indices analysis (parameters prioritization setting) θ_{mp} , θ_r , θ_m , K_1 , $k_{G,growth}$ and $k_{C,growth}$ were ranked as the six most influential parameters for the three phytoplankton groups, and also for nutrients state variables. K_{nit} is added to the list above accounting for nitrate concentration.

Global sensitivity analysis results have shown that the set of main parameters changes with time and it has a marked seasonal pattern because most of the main parameters are closely related to

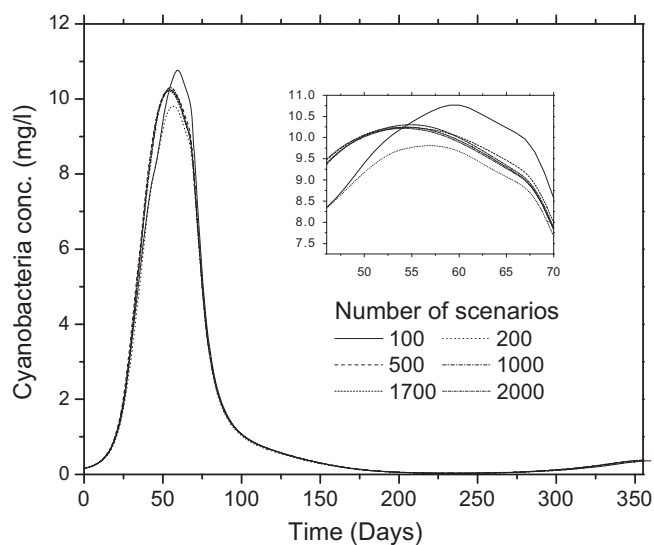


Fig. 1. Cyanobacteria mean concentration profiles as function of the number of scenarios.

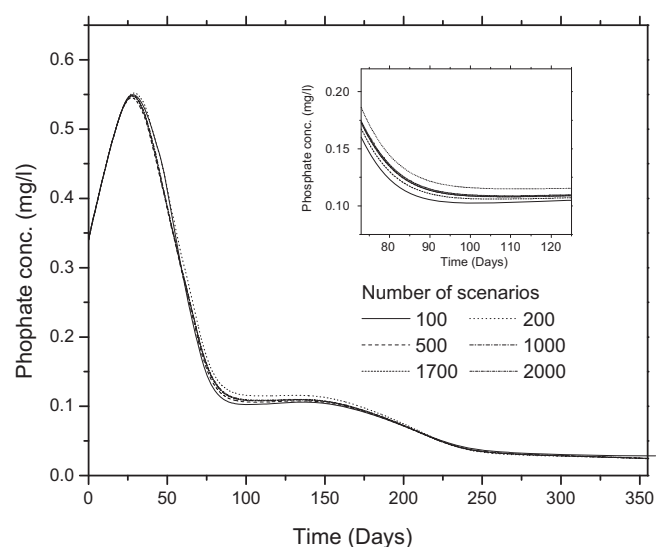


Fig. 2. Phosphate mean concentration profiles as function of the number of scenarios.

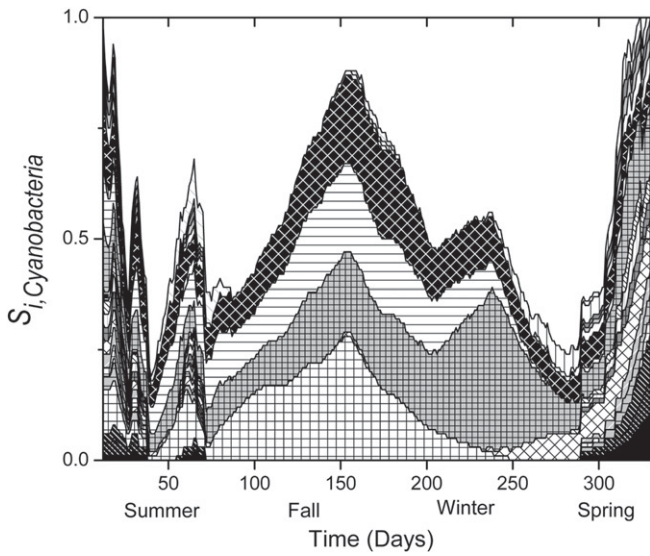


Fig. 3. S_i profiles for cyanobacteria concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{C,settling}$.

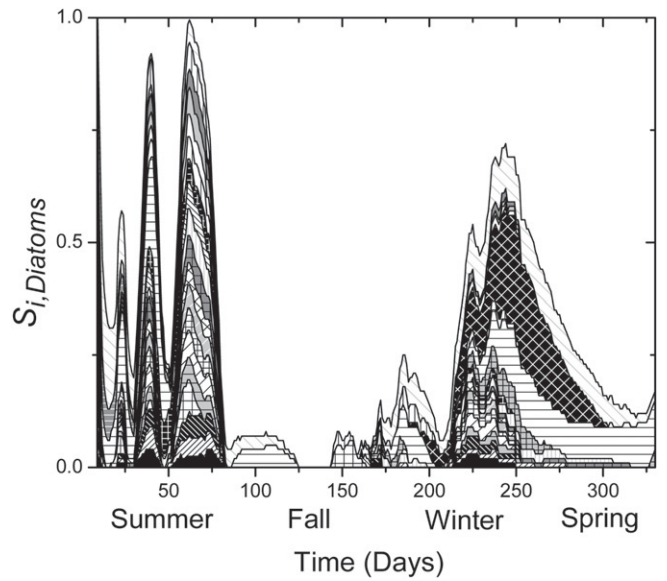


Fig. 5. S_i profiles for diatoms concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{C,settling}$.

the seasonal variables that drive the model, i.e., water temperature and solar radiation.

In the literature, background light attenuation, light limitation of phytoplankton growth related parameters and algae growth rates have been determined as main parameters in a number of freshwater and marine models with both local (Bierman and James, 1995; Omlin et al., 2001a; Chu et al., 2007; Fragoso et al., 2008) and global (Arhonditsis and Brett, 2005; Guven and Howard, 2007) sensitivity analysis. Results from the local sensitivity analysis performed by Omlin et al. (2001a) for a Lake Zürich biogeochemical, have shown that temperature dependence coefficient of biological processes contributed less to the uncertainty of model predictions than half-saturation light intensity of algae growth and algae

respiration, growth and death. In our study some of these coefficients (θ_{mp} , θ_r , θ_m) were found between the most influent parameters, both in first-order and total effects explaining large state variable variability (e.g. sum of θ_{mp} , θ_r , θ_m S_i for cyanobacteria concentration in day 150 is 57%). Guven and Howard (2007) applied a generalized sensitivity analysis to a model of cyanobacterial growth and found that the temperature dependence coefficient was the most critical parameter, while maximum growth rate and background light attenuation were the second and the last one in the list of 12 analyzed parameters. The parameters with stronger

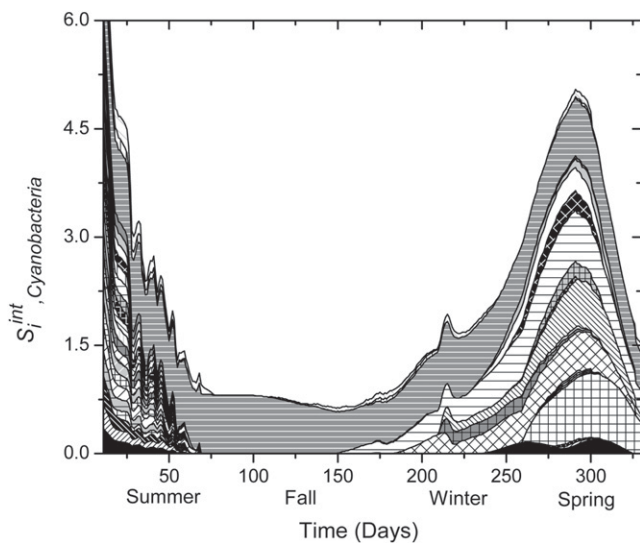


Fig. 4. S_i^{int} profiles for cyanobacteria concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{C,settling}$.

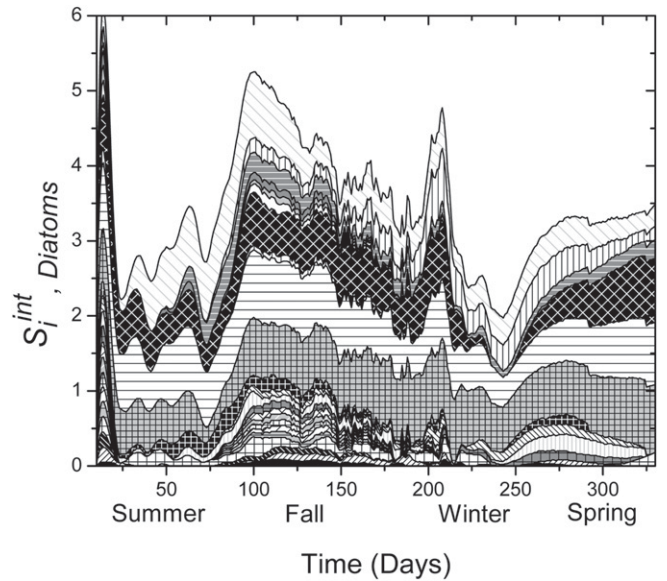


Fig. 6. S_i^{int} profiles for diatoms concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{C,settling}$.

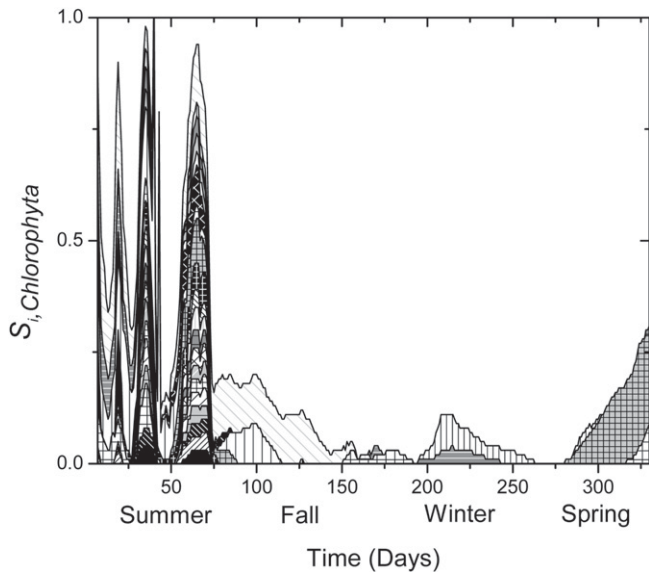


Fig. 7. S_i profiles for chlorophyta concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{G,settling}$.

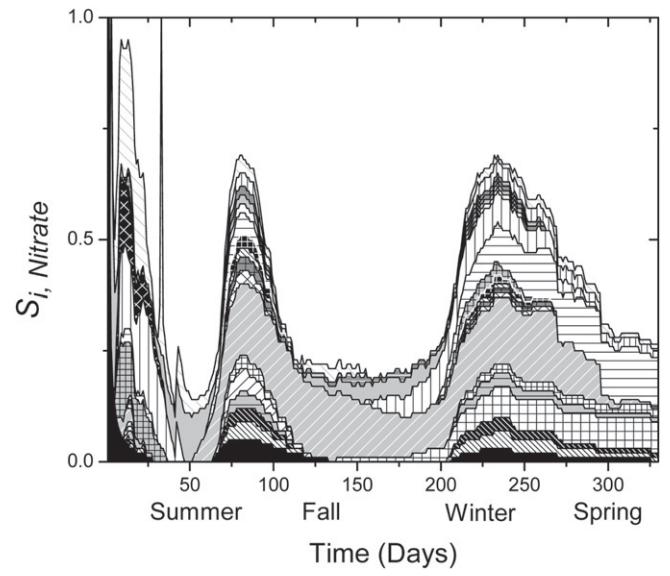


Fig. 9. S_i profiles for nitrate concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{G,settling}$.

effects for the model developed by Frago et al. (2008) were those related to the effects of temperature (temperature effect coefficient and respiration and excretion effect coefficient) and maximum growth rate of algae, but parameters associated to light penetration appeared to have a weaker effect on phytoplankton growth.

A sensitivity analysis based on Monte Carlo simulations and multivariate regression analysis for a eutrophication model of West Lake (China) showed that the settling rate of algae and temperature had the largest contribution to model prediction uncertainty of the state variables (Hongping and Yong, 2003). In our case, parameters associated to settling of the three phytoplankton groups ($k_{C,settling}$, $k_{D,settling}$, $k_{G,settling}$) modeled in for Paso de las Piedras Reservoir

were excluded of the calibration candidate list due to their low S_i^{TOT} values.

Cyanobacteria concentration is mainly driven by parameters related to meteorological conditions throughout the year and not by parameters related to other phytoplankton groups, both for first-order effects and interactions. Though $k_{C,growth}$ is not important in first effect terms, it is the main parameter that affects cyanobacteria concentration through interactions with other parameters. Chlorophyte and diatoms concentration profiles are more dominated by interactions between parameters than cyanobacteria through the year. These interactions, in addition to meteorological conditions, are related to other groups of phytoplankton and may be related to

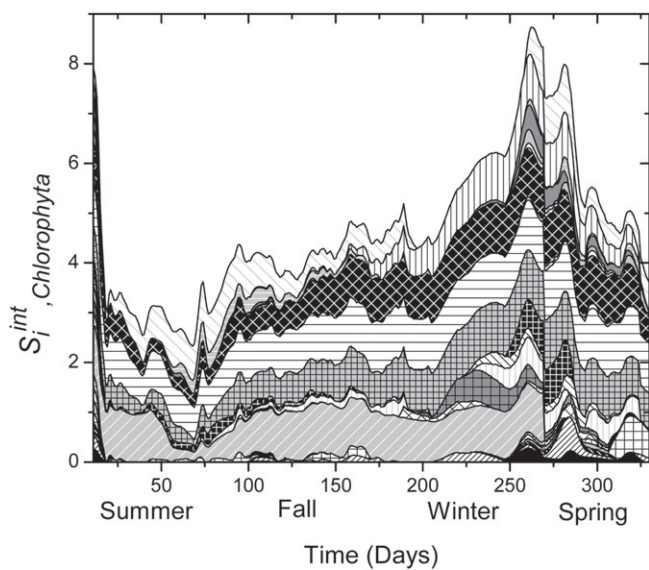


Fig. 8. S_i^{int} profiles for chlorophyta concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{G,settling}$.

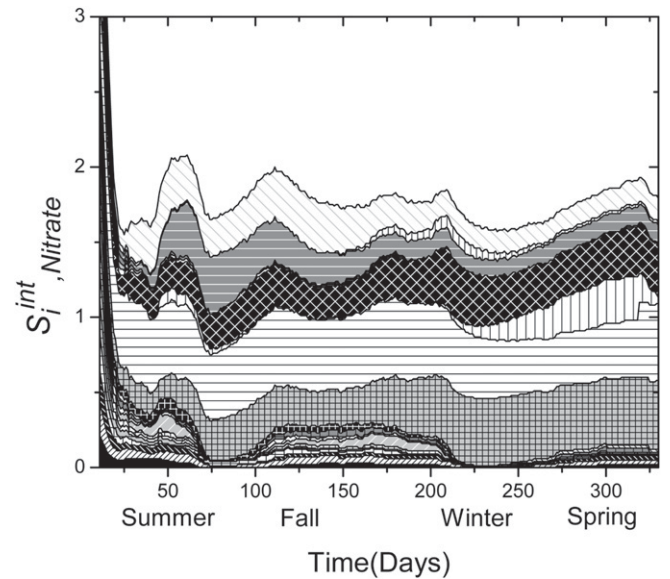


Fig. 10. S_i^{int} profiles for nitrate concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optc} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{G,settling}$.

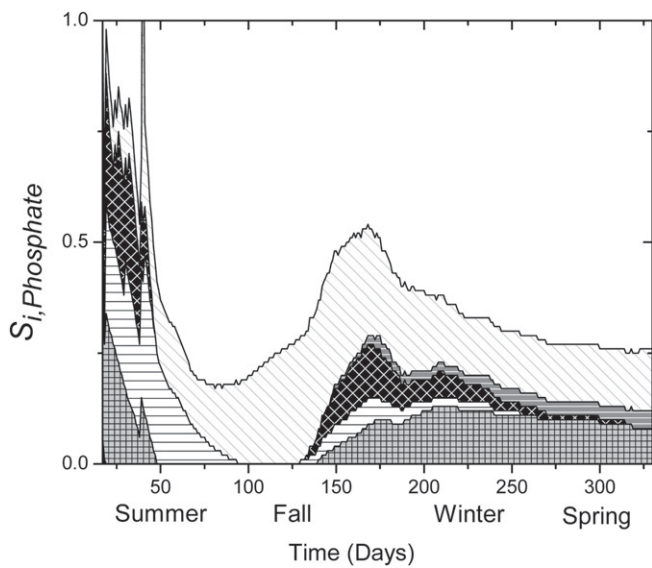


Fig. 11. S_i profiles for phosphate concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optC} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{G,settling}$.

the dominance of cyanobacteria and the competitive exclusion between ecological groups.

Even when sensitivity analysis results may be naturally dependent on model assumptions, local sensitivity analysis, which only takes into account infinitesimal variations of the parameter around the nominal value, assuming linear behavior, cannot quite capture model sensitivity to the parameter over its entire range of variation, especially in highly nonlinear models, as the eutrophication one described in this manuscript (Section 2.2). On the other global sensitivity analysis does not require the assumption of model linearity and/or additivity and explores the entire range of variation of the parameters, sampling from its probability distribution

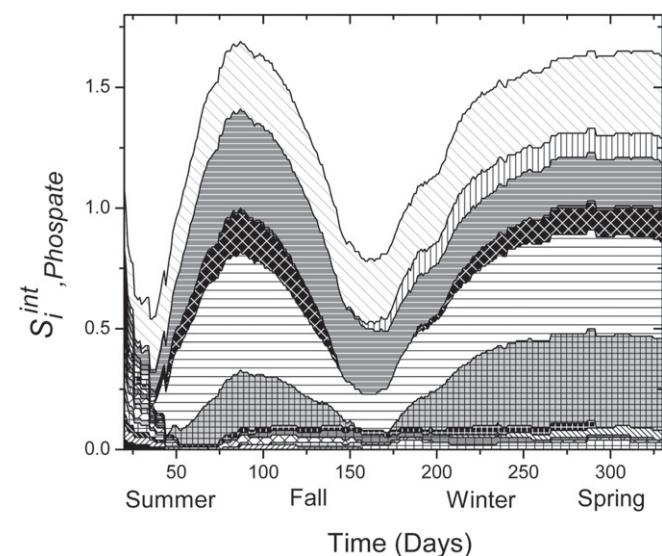


Fig. 12. S_i^{int} profiles for phosphate concentration. α_{nc} ; K_{mp} ; I_{optG} ; θ_{mn} ; $k_{C,growth}$; α_{pc} ; k_{nit} ; $k_{D,death}$; θ_{mp} ; $k_{D,growth}$; f_{ON} ; K_{nit} ; $k_{G,death}$; $k_{C,settling}$; $k_{G,growth}$; K_1 ; I_{optC} ; θ_m ; $k_{D,settling}$; K_{mn} ; I_{optD} ; θ_r ; $k_{G,settling}$.

Table 5

S_i and S_i^{int} for cyanobacteria concentration. Blanks correspond to zero values of sensitivity indices.

	S_i				S_i^{int}			
	$t = 34$	$t = 86$	$t = 180$	$t = 250$	$t = 34$	$t = 86$	$t = 180$	$t = 250$
$k_{C,grog}$					0.992	0.812	0.582	0.725
$k_{D,grog}$	0.013		0.006	0.033	0.092		0.048	
$k_{G,grog}$	0.017	0.008	0.010		0.142		0.069	
$k_{D,death}$	0.015				0.091			0.151
$k_{G,death}$	0.011				0.091			
$k_{C,settling}$	0.018		0.002		0.081			0.150
$k_{D,settling}$	0.015				0.091		0.011	0.092
$k_{G,settling}$	0.015				0.088			
I_{optC}	0.018				0.093			0.405
I_{optD}	0.015							0.178
I_{optG}	0.015				0.076			
k_{nit}	0.013				0.092			
k_{mn}	0.015							
k_{mp}					0.129			
K_{nit}	0.015				0.091			
θ_r	0.084	0.107	0.170	0.032			0.138	0.576
θ_m	0.076	0.085	0.181	0.269				
θ_{mn}	0.012				0.093			
θ_{mp}	0.091	0.078	0.194	0.074				
f_{ON}	0.013							
α_{nc}	0.013				0.092			0.092
α_{pc}	0.013				0.079			
K_1	0.026	0.100	0.148	0.003	0.054			
$\sum_{i=1}^{23} S_i$	0.529	0.378	0.711	0.437				

function. This feature justifies its higher computational cost against local sensitivity approaches. Due to the eutrophication model size, we do not perform different sensitivity analysis to compare with the global sensitivity analysis results reported in this work. However, for the sake of comparison of sensitivity analysis approaches performance, Cariboni et al. (2007) apply local and global strategies to ecological models, a classic Lotka–Volterra model and a fish population dynamics one. They conclude that the

Table 6

S_i and S_i^{int} for diatoms concentration. Blanks correspond to zero values of sensitivity indices.

	S_i				S_i^{int}			
	$t = 34$	$t = 86$	$t = 180$	$t = 250$	$t = 34$	$t = 86$	$t = 180$	$t = 250$
$k_{C,growth}$	0.013					0.417	0.174	0.097
$k_{D,growth}$	0.048					0.124		0.419
$k_{G,growth}$	0.023	0.037	0.061	0.102	0.480	0.867	0.541	0.391
$k_{D,death}$	0.030				0.052	0.282	0.033	0.146
$k_{G,death}$	0.005			0.002	0.052	0.282	0.033	
$k_{C,settling}$	0.020					0.051	0.013	
$k_{D,settling}$	0.021			0.010			0.031	
$k_{G,settling}$	0.024			0.008		0.005	0.013	
I_{optC}	0.019			0.007			0.028	
I_{optD}	0.032			0.011				0.074
I_{optG}			0.010	0.025	0.050	0.148	0.147	0.105
k_{nit}	0.023			0.005		0.031		
k_{mn}	0.020			0.008		0.031		
k_{mp}	0.027					0.031	0.009	
K_{nit}				0.008		0.031		
θ_r	0.086	0.004	0.045	0.128	0.793	0.788	0.604	0.390
θ_m			0.022	0.057	0.616	0.379	0.824	0.605
θ_{mn}	0.028			0.008		0.035		
θ_{mp}				0.241	0.629	0.387	0.479	0.038
f_{ON}	0.023			0.005		0.031		
α_{nc}	0.023			0.005				0.031
α_{pc}	0.018			0.009		0.009	0.023	
K_1					0.179	0.115	0.016	
$\sum_{i=1}^{23} S_i$	0.501	0.042	0.129	0.652				

(2001). As it is reported in Estrada et al. (2009a), after parameter estimation for the fifteen most important parameters, average deviations between observed data and model predictions for concentrations of cyanobacteria, phosphate and nitrate are 0.11; 0.04 and 0.12, respectively. The model with the adjusted set of parameters closely reproduces cyanobacteria autumn bloom, as well as summer bloom, mainly composed of chlorophytes and cyanobacteria (80 and 20%, respectively) and spring blooms of diatoms in the lake.

5. Conclusions

In this paper we have applied a global variance-based sensitivity analysis to a highly nonlinear first principle-based dynamic eutrophication model. Global sensitivity analysis is a model independent technique, based on estimating the fractional contribution of each parameter to the variance of the model variables. It also takes into account interactions between parameters and provides qualitative accurate sensitivity measures. By applying Sobol's–Saltelli's algorithm, we have calculated time profiles for first-order and total sensitivity indices to determine non-influential parameters and select the set for calibration purposes according to their importance. Global sensitivity analysis results have shown that the set of main parameters changes with time and it has a marked seasonal pattern, as expected, because most of the control parameters are closely related to the seasonal variables that drive the model, i.e., water temperature and solar radiation (θ_{mp} , θ_p , θ_m and K_1). Since the summation of the first-order sensitivity indices is most of the time lower than 1, we can conclude that the model is not simply additive in model parameters, mainly regarding diatom and chlorophyte concentrations. Based on parameter ranking by global sensitivity analysis, we have estimated these parameters with advanced dynamic optimization techniques and available observed data from Paso de las Piedras reservoir (Estrada et al., 2009a).

References

- Anh, D.T., Bonnet, M.P., Vachaud, G., Van Minh, C., Prieur, N., Vu Duc, L., Anh, L.L., 2006. Biochemical modeling of the Nhue River (Hanoi, Vietnam): practical identifiability analysis and parameters estimation. *Ecological Modelling* 193, 182–193.
- Arhonditsis, G.B., Brett, M.T., 2005. Eutrophication model for Lake Washington (USA) Part I. Model description and sensitivity analysis. *Ecological Modelling* 187, 140–178.
- Arhonditsis, G.B., Brett, M.T., 2005a. Eutrophication model for Lake Washington (USA) Part I. Model description and sensitivity analysis. *Ecological Modelling* 187, 140–178.
- Arhonditsis, G.B., Brett, M.T., 2005b. Eutrophication model for Lake Washington (USA) Part II. Model calibration and system dynamic analysis. *Ecological Modelling* 187, 179–200.
- Baklouti, M., Faure, V., Pawlowski, L., Sciandra, A., 2006. Investigation and sensitivity analysis of mechanistic phytoplankton model implemented in a new modular numerical tool (Eco3M) dedicated to biogeochemical modelling. *Progress in Oceanography* 71, 34–58.
- Bard, Y., 1974. *Nonlinear Parameter Estimation*. Academic Press, Cambridge, MA.
- Bierman, V.J., James, R.T., 1995. A preliminary modeling analysis of water quality in Lake Okeechobee, Florida: diagnostic and sensitivity analyses. *Water Research* 29, 2767–2775.
- Brun, R., Reichert, P., Kúnsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. *Water Resources Research* 37, 1015–1030.
- Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software* 22, 1509–1518.
- Cariboni, J., Gatelli, D., Liska, R., Saltelli, A., 2007. The role of sensitivity analysis in ecological modelling. *Ecological Modelling* 203, 167–182.
- Chu, P.C., Ivanov, L.M., Margolina, T.M., 2007. On non-linear sensitivity of marine biological models to parameter variations. *Ecological Modelling* 206, 369–382.
- Cossarini, G., Solidoro, C., 2008. Global sensitivity analysis of a trophodynamic model of the Gulf of Trieste. *Ecological Modelling* 212, 16–27.
- Cukier, R.I., Fortuin, C.M., Shuler, K.E., Petschek, A.G., Schaibly, J.H., 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. *Journal of Chemical Physics* 59, 3873–3878.
- Cukier, R.I., Levine, H.B., Shuler, K.E., 1978. Nonlinear sensitivity analysis of multi-parameter model systems. *Journal of Computational Physics* 26, 1–42.
- Estrada, V., Parodi, E.R., Diaz, M.S., 2009a. Determination of biogeochemical parameters in eutrophication models as large scale dynamic parameter estimation problems. *Computers and Chemical Engineering* 33, 1760–1769.
- Estrada, V., Parodi, E.R., Diaz, M.S., 2009b. Addressing the control problem of algae growth in water reservoirs with advanced dynamic optimization approaches. *Computers & Chemical Engineering* 33, 2063–2074.
- Fragoso Jr., C.R., Motta Marques, D.M.L., Collischonn, W., Tucci, C.E.M., van Nes, E.H., 2008. Modelling spatial heterogeneity of phytoplankton in Lake Mangueira, a large shallow subtropical lake in South Brazil. *Ecological Modelling* 219, 125–137.
- Güven, B., Howard, A., 2007. Identifying the critical parameters of a cyanobacterial growth and movement model by using generalised sensitivity analysis. *Ecological Modelling* 207, 11–21.
- Häkanson, L., 2000. The role of characteristic coefficients of variation in uncertainty and sensitivity analyses, with examples related to the structuring of lake eutrophication models. *Ecological Modelling* 131, 1–20.
- Hamilton, D.P., Schladow, S.G., 1997. Prediction of water quality in lakes and reservoirs. Part I. Model description. *Ecological Modelling* 96, 91–110.
- Hangos, K., Cameron, I., 2001. *Process Modelling and Model Analysis*. Academic Press.
- Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety* 52, 1–17.
- Hongping, P., Yong, W., 2003. Eutrophication research of West Lake, Hangzhou, China: modeling under uncertainty.
- Intartaglia, C., Sala, S.E., 1989. Variación estacional del fitoplancton en un lago no estratificado: Embalse Paso de las Piedras, Argentina. *Revista Brasileira de Biología* 49, 873–882.
- Koda, M., McRae, G.J., Seinfeld, J.H., 1979a. Automatic sensitivity analysis of kinetic mechanisms. *International Journal of Chemical Kinetics* 11, 427–444.
- Koda, M., Dogru, A.H., Seinfeld, J.H., 1979b. Sensitivity analysis of partial differential equations with application to reaction and diffusion processes. *Journal of Computational Physics* 30, 259–282.
- Liepmann, D., Stephanopoulos, G., 1985. A dynamic model of a closed ecosystem: development and global sensitivity analysis. *Ecological Modelling* 30, 13–30.
- Lindenschmidt, K.E., 2006. The effect of complexity on parameter sensitivity and model uncertainty in river water quality modelling. *Ecological Modelling* 190, 72–86.
- Manache, G., Melching, C.S., 2008. Identification of reliable regression- and correlation-based sensitivity measures for importance ranking of water-quality model parameters. *Environmental Modelling & Software* 23, 549–562.
- McRae, G.J., Tilden, J.W., Seinfeld, J.H., 1982. Global sensitivity analysis—a computational implementation of the Fourier Amplitude Sensitivity Test (FAST). *Computers & Chemical Engineering* 6, 15–25.
- Mieleitner, J., Reichert, P., 2006. Analysis of the transferability of a biogeochemical lake model to lakes of different trophic state. *Ecological Modelling* 194, 49–61.
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* 33, 161–174.
- Nguyen, T.G., Kok, J.L., 2007. Systematic testing of an integrated systems model for coastal zone management using sensitivity and uncertainty analyses. *Environmental Modelling & Software* 22, 1572–1587.
- Norton, J.P., 2008. Algebraic sensitivity analysis of environmental models. *Environmental Modelling & Software* 23, 963–972.
- Omlin, M., Brun, R., Reichert, P., 2001a. Biogeochemical model of Lake Zürich: model equations and results. *Ecological Modelling* 141, 77–103.
- Omlin, M., Reichert, P., Forster, R., 2001b. Biogeochemical model of Lake Zürich: sensitivity, identifiability and uncertainty analysis. *Ecological Modelling* 141, 105–123.
- Parodi, E., Estrada, V., Trobbiani, N., Argañaraz Bonini, G., 2004. Análisis del estado trófico del embalse Paso de las Piedras. *Ecología en tiempos de Cambio*, 178.
- Pastres, R., Chart, K., Solidoro, C., Dejak, C., 1999. Global sensitivity analysis of a shallow-water 3D eutrophication model. *Computer Physics Communications* 117, 62–74.
- Pastres, R., Ciavatta, S., Cossarini, G., Solidoro, C., 2003. Sensitivity analysis as a tool for the implementation of a water quality regulation based on the maximum permissible loads policy. *Reliability Engineering and System Safety* 79, 239–244.
- Pastres, R., Ciavatta, S., 2005. A comparison between the uncertainties in model parameters and in forcing functions: its application to a 3D water-quality model. *Environmental Modelling & Software* 20, 981–989.
- Process Systems Enterprise Limited, 2009. *gProms Model Developer Guide* (www.pcenterprise.com).
- Saloranta, T.M., Andersen, T., 2007. MyLake – a multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. *Ecological Modelling* 207, 45–60.
- Saltelli, A., Sobol', I.M., 1995. About the use of rank transformation in sensitivity analysis of model output. *Reliability Engineering and System Safety* 50, 225–239.
- Saltelli, A., Tarantola, S., Chan, K., 1999. Quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* 41, 39–56.
- Saltelli, A., Tarantola, S., 2002. On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal. *Journal of the American Statistical Association* 97, 702–709.

- Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity analysis in practice. A guide to assessing scientific models. In: Probability and Statistics Series. John Wiley & Sons Publishers.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis. The Primer. John Wiley & Sons Publishers.
- Schefer, J.C. 2004. Los recursos hídricos y el abastecimiento de agua. Región de Bahía Blanca. CEPADE, Argentina. 132–137.
- Schladow, S.G., Hamilton, D.P., 1997. Prediction of water quality in lakes and reservoirs. Part 2. Model calibration, sensitivity analysis and application. *Ecological Modelling* 96, 111–123.
- Sobol', I.M., 1990. Sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie* 2, 112–118 [in Russian]. Translated in English in Sobol', 1993. *Mathematical Modelling and Computational Experiment* 1, 407–414.
- Sobol', I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 55, 271–280.
- Sobol', I.M., Levitan, Y.L., 1999. On the use of variance reducing multipliers in Monte Carlo computations of a global sensitivity index. *Computer Physics Communications* 117, 52–61.
- Zhang, J., Jørgensen, S.E., Mahler, H., 2004. Examination of structurally dynamic eutrophication model. *Ecological Modelling* 173, 313–333.