# Water quality modelling for small river basins

Available online at www.sciencedirect.com
ScienceDirect
Environmental Modelling & Software 23 (2008) 451–463
Environmental Modelling & Software
www.elsevier.com/locate/envsoft

# Water quality modelling for small river basins

Stefano Marsili-Libelli*, Elisabetta Giusti

*Department of Systems and Computers, University of Florence, Via S. Marta, 3, 50139 Firenze, Italy*

Received 2 September 2006; received in revised form 19 June 2007; accepted 20 June 2007
Available online 13 August 2007

**Abstract**

Water quality modelling in small rivers is often considered unworthy from a practical and economic viewpoint. This paper shows instead that a simple model structure can be set up to describe the stationary water quality in small river basins in terms of carbon and nitrogen compounds, when the use of complex models is unfeasible. In short rivers point and nonpoint sources play a key role in shaping the model response, being as important as the self-purification dynamics. Further, the varying river characteristics, in terms of morphology, hydraulics and vegetation, require the introduction of variable parameters, thus complicating the originally simple model structure. To determine the identifiability of the resulting model an identifiability assessment was carried out, based on sensitivity analysis and optimal experiment design criteria. The identifiable subset was determined by ranking the parameters in terms of sensitivity and computing the associated Fisher Information Matrices. It was found that the inclusion of the nonpoint sources as piecewise constant parameters affected the identifiability to a considerable extent. However, the combined parameter–sources calibration was made possible by the use of a robust estimation algorithm, which also provided estimation confidence limits. The calibrated model responses are in good agreement with the data and can be used as scenario generators in a general strategy to conserve or improve the water quality.


*Keywords:* River quality; Model identification; Parameter estimation; Ecological models; Sensitivity; Uncertainty analysis

## 1. Introduction

In developing water quality models small river basins pose specific problems due to data scarcity, lack of major investments as a consequence of their minor importance, and the large number of diverse inputs, especially if they flow through densely populated areas. For these reasons it is difficult to adapt major water quality models, such as those provided by the US Environmental Protection Agency: QUAL2E (Brown and Barnwell, 1987), QUAL2K (Chapra and Pellettier, 2003), WASP6 (Wool et al., 2006), or the IWA River Quality Model No. 1 (Reichert et al., 2001), which require more information regarding the river system than is often available. In these cases, it makes sense to use *ad hoc* simple models in order to derive the crucial information about the river quality and become part of a decision support system.

The idea of using model structures that are appropriate for the application and available data was first advocated by Halfon (1983) and Beck (1987). Pearl (1978) has also noted that the tendency to prefer simple models is partly based on the notion that they are easier to calibrate and therefore more reliable, as confirmed by Snowling and Kramer (2001), who showed that complex model is generally very sensitive and therefore difficult to identify, and Lindenschmidt (2006), who pointed out that the most complex model is not necessarily the most accurate. On the other hand data scarcity may lead to heavily overdetermined calibration problems, because the large number of adjustable parameters and the few experimental constraints make good fits almost guaranteed, at least using mean parameter values. To investigate potentially large errors on output created by interacting errors on the parameters, a multiple sensitivity analysis (Haefner, 2005) was performed by comparing the effect of independent parameter perturbations on the model outputs. Each parameter value was drawn from an independent Gaussian pdf with variance equal to the

* Corresponding author. Tel./fax: +39 055 47 96 264.
*E-mail address:* marsili@dsi.unifi.it (S. Marsili-Libelli).

1364-8152/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2007.06.008

**Nomenclature**

$B_d$ (mg $O_2$ $L^{-1}$ $km^{-1}$) nonpoint source CBOD load
$C_t$ ($°C^{-1}$) temperature coefficient
$DO_{ph}$ (mg $O_2$ $L^{-1}$ $km^{-1}$) photosynthetic oxygen production
$\boldsymbol{F}$ Fisher Information Matrix ($n_p \times n_p$)
$K_{aa}$ ($day^{-1}$) first order $NH_4$-N decay rate
$K_{al}$ ($day^{-1}$) nitrogen algal uptake maximum rate
$K_{a\_max}$ ($day^{-1}$) nitrification maximum rate
$K_b$ ($day^{-1}$) CBOD decay rate
$K_{bm}$ ($day^{-1}$) average CBOD decay coefficient
$K_c$ ($day^{-1}$) reaeration coefficient
$K_f$ (mg N $L^{-1}$) half-saturation constant for N uptake by algae
$K_o$ ($day^{-1}$) denitrification rate
$K_{os}$ (mg $O_2$ $L^{-1}$) half-saturation constant for $NH_4$-N oxidation rate
$K_{sa}$ (mg N-$NO_3$ $L^{-1}$) nitrification half-saturation constant (Michaelis–Menten)
$K_r$ ($day^{-1}$) reaeration rate
$k$ index denoting the sampling locations along the river length
$N$ number of experimental data
$NO_{3d}$ (mg $NO_3$-N $L^{-1}$ $km^{-1}$) nonpoint source nitrate-N load
$\boldsymbol{Q}_k$ matrix of indexed weighting factors related to the measurement noise
$S_{i,j} = \partial y_j/\partial p_i$ sensitivity of the $j$-th output to the $i$-th parameter
$\boldsymbol{p}$ vector of $n_p$ model parameters
p$K$ pH coefficient
$t_{N-1}^{\alpha/2}$ two-tailed student-$t$ statistics at 100 $(1-\alpha)$% confidence level for $N-1$ degrees of freedom
$t_s$ flow time
$u$ (m $s^{-1}$) stream velocity
$Y_a$ Autotrophic organism yield factor
$\boldsymbol{y}_{j,k}^{exp}$ indexed noisy experimental observations ($j = 1, \ldots, q$)
$\boldsymbol{y}_{j,k}$ indexed parameter-dependent model outputs ($j = 1, \ldots, q$)

*Acronyms*
CBOD Carbonaceous Biochemical Oxygen Demand
NBOD Nitrogenous Biochemical Oxygen Demand
OED Optimal Experiment Design
PDE Partial Differential Equation
pdf Probabilitiy densitiy function
PEAS Parameter Estimation and Accuracy Software
WWTP Waste Water Treatment Plant

*Greek letters*
$\delta$ preferential coefficient for $NH_4$-N uptake by algae
$\delta p_i$ parameter confidence limits
$\xi$ global sensitivity index

squared estimated confidence interval. The resulting simulations were then compared to the uncertainty ranges of the data, in order to quantify the possible spread of model responses.

Given the large number of model parameters vis-à-vis the limited number of data, a minimal identifiable subset had to be identified. This was obtained using Optimal Experiments Design (OED) criteria (Fedorov, 1972; Atkinson and Donev, 1992) based on the Fisher Information Matrix, which yields an indication of the information content of a given experiment (model plus data).

This paper presents a simple water quality model to describe the dynamics of carbon and nitrogen in small river basins together with a robust estimation scheme (Marsili-Libelli et al., 2003; Checchi and Marsili-Libelli, 2005; Marsili-Libelli and Checchi, 2005; Checchi et al., 2007) for calibrating its parameters. To take into account the varying nature of the river along its course, some parameter variability was introduced through switching functions and the identifiability implications have been analyzed. Two minor but environmentally important rivers in Tuscany (central Italy) were considered: the Sieve and the Ombrone. They are both important tributaries to the river Arno, as shown in Fig. 1, and play a key role in determining the water quality in the whole river system.

The paper is organized as follows: first the basic model structure is introduced, the identifiability method is described and then parameter estimation is considered. The validity of the calibrated models and their use as a scenario generation tool is then discussed.

### 1.1. Basins' description

The two rivers considered in this study, Sieve and Ombrone, are both part of the Arno river system and have widely differing characteristics. Hydraulically both rivers are subjected to massive winter and autumn floods, but reduce to almost a trickle during the summer months, when the river quality becomes critical and it is precisely then that a water quality model is needed. The models were developed for typical low-flow conditions, which were determined by averaging the daily flows during the summer months to obtain a typical flow of 2.9 $m^3$ $s^{-1}$ for the Sieve and 1.7 $m^3$ $s^{-1}$ for the Ombrone.

#### 1.1.1. Sieve basin

The quality of the Sieve river is definitely the best in the Arno basin and a great deal of efforts is being made to conserve its pristine conditions. The reach considered here extends from the upstream Bilancino reservoir down to the confluence with the Arno, for a total length of 47 km. The water released from the Bilancino reservoir, normally 3 $m^3$ $s^{-1}$, is intended to guarantee the minimum sustainable flow during the summer months and feeds enough water to the Florence potabilization plant, some 20 km downstream of the confluence into the Arno.

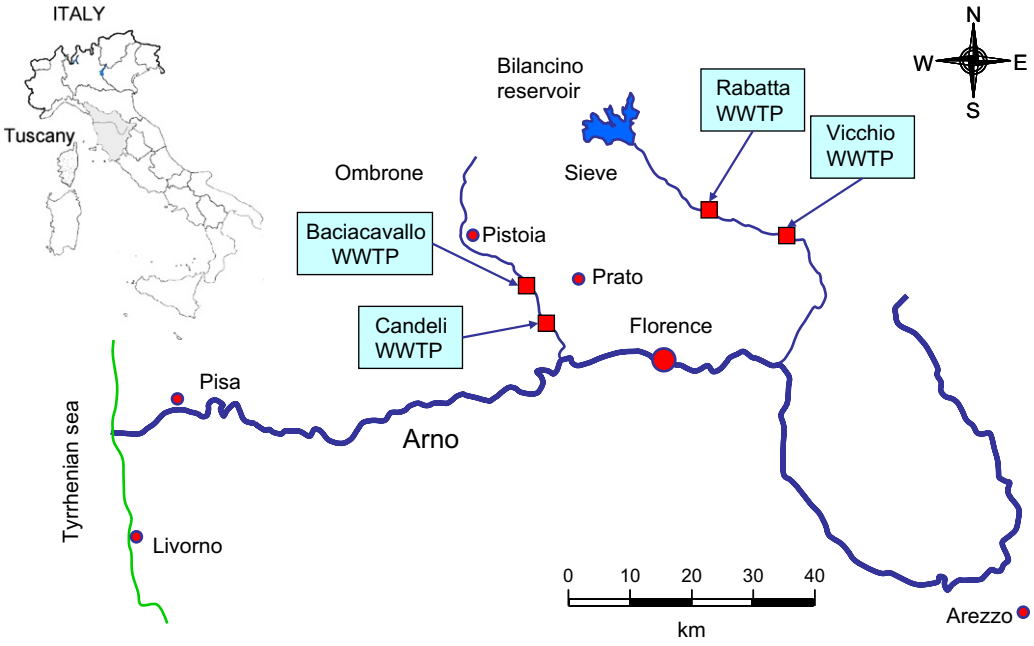


Fig. 1. Location of the Ombrone and Sieve rivers, representing two of the main right bank tributaries to the Arno river, in Tuscany (shaded area in the upper-left map).

The Sieve catchment is rather large and includes many tributaries which may increase the downstream flow by more than fourfold with respect to the upstream value, as shown in Fig. 2. The slope of the reach is less pronounced in the upstream part, up to Vicchio, and increases in the final part. This, together with the major tributaries located in the downstream part, makes the final part of the river much richer in flow and speed than the upstream part, contrary to what happens in most rivers where the upstream part is steeper and fast-flowing. However, the increased depth in this part counterbalances the larger flow and no increase in reaeration rate was observed. There are two wastewater treatment plants discharging in the Sieve river, located at Rabatta and downstream of Vicchio, whereas many untreated discharges and nonpoint source pollution from agriculture, detected during this study, constitute a considerable contribution to the river pollution. In setting up the model, information about these sources was obtained through direct interview with local municipal officers or was estimated in the model calibration phase, as explained later. The river description in terms of characteristics and sources (point and nonpoint) is shown in Fig. 3.

### *1.1.2. Ombrone basin*

Contrary to the Sieve, the Ombrone river flows through a heavily industrialized area dominated by the textile centre of Prato and is subject to many discharges, some of which

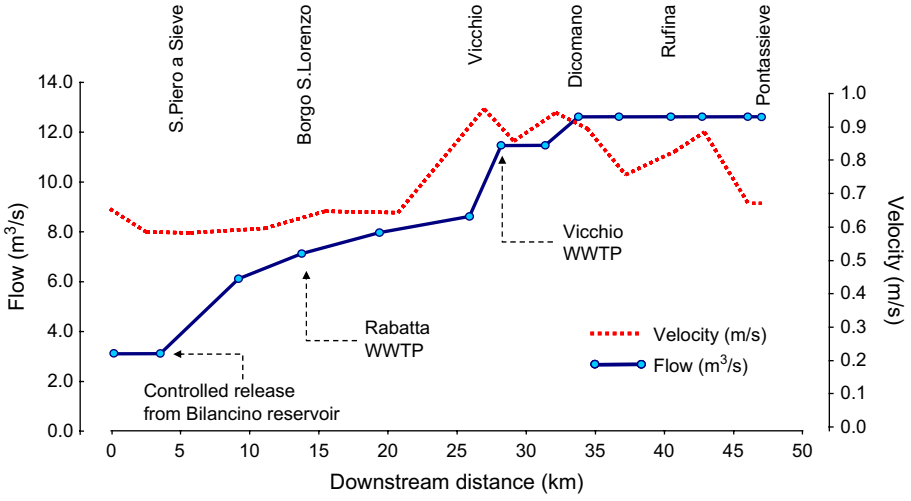


Fig. 2. Flow and velocity reconstruction along the Sieve river used in the water quality model. The upstream flow is controlled by the Bilancino reservoir and the many tributaries produce a fourfold flow increase at the end of the reach.

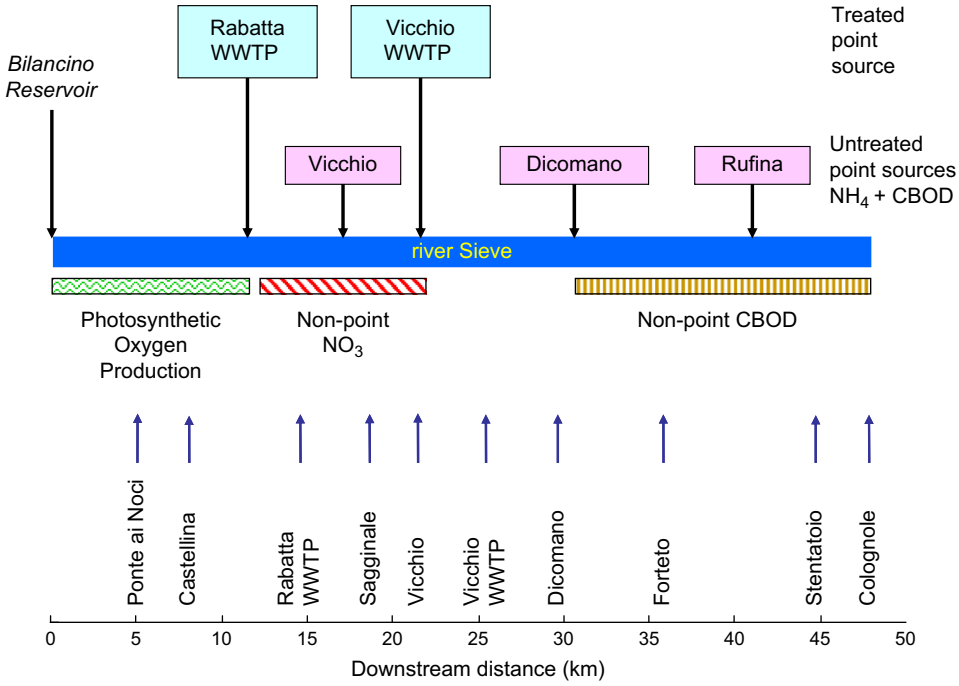


Fig. 3. Distribution of measuring points (upward arrows), point and nonpoint sources along the Sieve river.

untreated, in addition to the output of the major wastewater treatment works of Baciacavallo, with a capacity exceeding 750,000 PE. A smaller wastewater treatment plant is located at Candeli, 3 km downstream of Baciacavallo, with a capacity of 5000 PE. The reach under study is 15 km in length and can be divided into two parts with widely differing characteristics. The upstream part, 10 km long, flows through a densely populated area, with artificial riverbanks and many pollution sources, whereas the downstream reach, of about 5 km, is relatively unspoiled with densely vegetated riparian zones and sparse human settlements. In this reach the photosynthetic activity is considerable, given the high nutrient content of the incoming water, and self-purification is very active. The river flow and velocity are shown in Fig. 4, whereas the loading and sampling details are shown in Fig. 5.

### 1.2. Materials and methods

The water quality measurements were taken directly on site using a OXI90 (WTW – Wissenschaftlich-Technische Werkstätten GmbH, Weilheim, Germany) for dissolved oxygen and a Spectroquant NOVA60® Photometer (Merck KGaA, Darmstadt, Germany) for COD and inorganic nitrogen compounds, using test kits COD 160 C1/25, 15–160 mg $L^{-1}$ COD (method 013); ammonium-N A5/25, 0.20–8.00 mg $L^{-1}$ $NH_4$-N indophenol (method 048); and nitrate-N 50 N1/25, 0.5–23.0 mg

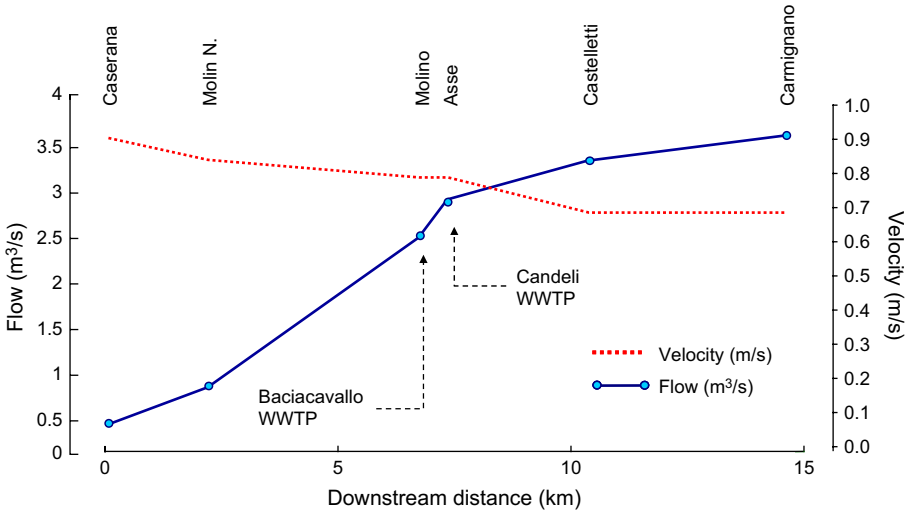


Fig. 4. Flow and velocity reconstruction along the Ombrone river used in the water quality model. The Baciacavallo WWTP discharging near the Molino sampling point significantly contributes to the downstream flow, especially during the low-flow season.

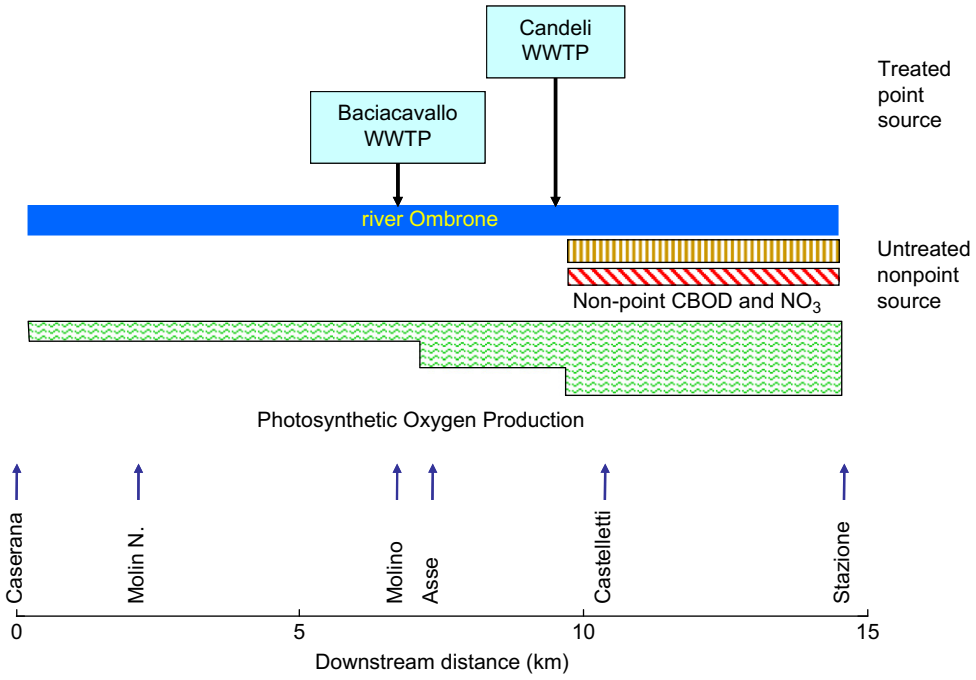


Fig. 5. Distribution of measuring points (upward arrows), point and nonpoint sources along the Ombrone river.

$L^{-1}$ $NO_3$-N 2,6-dimethylphenol (method 029). For the Ombrone river the cooperation of the provincial laboratory of the Regional Environmental Protection Agency (ARPAT) is gratefully acknowledged for sampling and water quality analysis. The sampling points were georeferenced using a Garmin GPS12 (Garmin, Olathe, KS, USA) portable global positioning system. From the accuracy viewpoint, the data values and associated error bars appearing in the model response figures were computed as the mean and standard deviation of the samples taken in the same transect on differing dates between July and October 2001 in the river Sieve and in the same period of 2005 in the Ombrone.

## 2. Model structure

### 2.1. Basic modelling assumptions

The general one-dimensional advective–diffusive dynamics for a reactive pollutant can be written as a PDE (see e.g. Rinaldi et al., 1979; Thomann and Mueller, 1987; Chapra, 1997)

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} + D\frac{\partial^2 C}{\partial x^2} - f(C), \tag{1}$$

where $C$ (mg $L^{-1}$) is the concentration of a generic pollutant, $u$ is the stream velocity (m $s^{-1}$), $D$ is the diffusion coefficient ($m^2$ $s^{-1}$), $t$ is the time (s), $x$ is the downstream distance along the river length (m) and $f(C)$ is the generic reactive term for the pollutant $C$. Neglecting the diffusion term, Eq. (1) becomes

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} - f(C). \tag{2}$$

Taking the full differential of $C$ yields

$$\mathrm{d}C = \frac{\partial C}{\partial x}\mathrm{d}x + \frac{\partial C}{\partial t}\mathrm{d}t. \tag{3}$$

Dividing by d$t$ and taking into account Eq. (2) yields

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \frac{\partial C}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial C}{\partial t} = \frac{\partial C}{\partial x}u + \frac{\partial C}{\partial t} = -f(C). \tag{4}$$

Therefore, defining the flow time $t_s$ as $\mathrm{d}t/\mathrm{d}t_s = 1$ and $\mathrm{d}x/\mathrm{d}t_s = u$, the PDE (Eq. (2)) becomes a simple ordinary differential equation (Eq. (4)), whose solution is valid not on the entire ($t$, $x$) domain, but only on the characteristic line defined by $u$. In this case, the model Eq. (4) can be expressed in flow time $t_s$, i.e. the travel time of a given plug of water along the river reduces to the reactive term $f(C)$. To remain on the characteristic line the measurements must be taken using the "follow-the-plug" assumption, where $\mathrm{d}x/\mathrm{d}t_s = u$ and Eq. (4) fully describes the river quality behaviour.

### 2.2. Introduction of reactive terms

The reactive processes included in the model are: degradation of dissolved carbonaceous substances, ammonium oxidation, algal uptake and denitrification, and dissolved oxygen balance, including depletion by degradation processes and supply by physical reaeration and photosynthetic production. The interconnections among these processes, together with the nonpoint source inputs, are shown in Fig. 6 and the model equations are listed in Table 1. Since the model is space-referenced through the "follow-the-plug" assumption, the variation

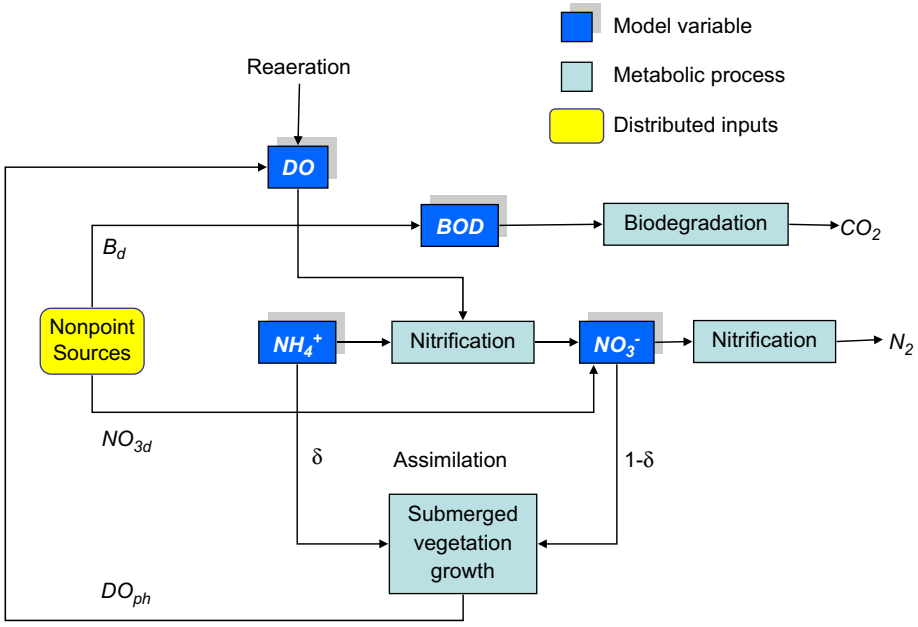


Fig. 6. Transformation processes included in the water quality model.

of flow and velocity along the course was included in the model, as already discussed in Section 1.1. Given the differing morphology of the two rivers, the general model of Table 1 had to be adapted to each river including the appropriate processes as shown in Table 2.

To compare this model structure with the complexity grades introduced by Lindenschmidt (2006), the BOD was differentiated between carbonaceous (CBOD) and nitrogenous components (NBOD) (complexity 2) with a loose coupling between the DO–BOD cycle and the algae–nitrogen interaction. Phosphorus was not included in the model because in both rivers its concentration was too low to be observed and modelled reliably. Further, almost the entire algal population is composed of N-limited species and its interaction with dissolved inorganic nitrogen is described by the preferential absorption coefficient $\delta$ introduced by Brown and Barnwell (1987) in the QUAL2 model family. The relatively simple model structure is partially offset by structuring some parameter as a function of the varying river morphology. In fact it was observed that only relating some parameters to the river

Table 1
Structure of the river water quality model

Dynamics

$$\frac{dB}{dx} = -\frac{1}{u(x)}K_b(x)B + B_d(x)$$

$$\frac{d\mathrm{NH_4}}{dx} = -\frac{1}{u(x)}K_a(x)\mathrm{NH_4} - \frac{1}{u(x)}\delta K_{al}\frac{\mathrm{NH_4}}{K_f + \mathrm{NH_4}}$$

$$\frac{d\mathrm{NO_3}}{dx} = \frac{1}{u(x)}K_a(x)\mathrm{NH_4} - \frac{1}{u(x)}(1-\delta)K_{al}\frac{\mathrm{NO_3}}{K_f + \mathrm{NO_3}} - \frac{1}{u(x)}K_o\mathrm{NO_3} + \mathrm{NO_{3d}}(x)$$

$$\frac{d\mathrm{DO}}{dx} = \frac{1}{u(x)}K_r(u)(\mathrm{DO_{sat}} - \mathrm{DO}) - \frac{1}{u(x)}K_b(x)B - \frac{1}{u(x)}\frac{(4.57 - Y_a)}{Y_a}K_a(x)\mathrm{NH_4} + \mathrm{DO_{ph}}(x)$$

Sieve

$$K_a(x) = \begin{cases} K_{a\_max}\frac{\mathrm{NH_4}}{K_{sa}+\mathrm{NH_4}} & x > 21 \\ K_{aa} & x < 21 \end{cases}$$

$$K_b(x) = K_{bm}\begin{cases} 1.802 & 0 < x \le 12.492 \\ 2.479 & 12.492 < x < 16.859 \\ 0.451 & 16.859 < x < 48.454 \end{cases}$$

Ombrone

$$K_a(x) = \begin{cases} K_{a\_max1}f(\mathrm{T,pH,DO}) & 0 < x < 10.5 \\ K_{a\_max2}f(\mathrm{T,pH,DO}) & x > 10.5 \end{cases}$$

$$f(\mathrm{T,pH,DO}) = e^{C_T(\mathrm{T}(x)-15)}\frac{1}{1 + 10^{pk_1 - \mathrm{pH}(x)} + 10^{\mathrm{pH}(x) - pK_2}}\left(\frac{\mathrm{DO}}{K_{os} + \mathrm{DO}}\right)$$

$$K_b(x) = \begin{cases} K_{b1} & 0 < x < 10.5 \\ K_{b2} & x > 10.5 \end{cases}$$

$$K_r(u) = K_c\sqrt{\frac{u(x)}{3.6}}$$

Table 2
Processes included in each model (denoted by ×)

| Process | Sieve | Ombrone |
|---|---|---|
| BOD degradation | × | × |
| BOD nonpoint source | × | |
| $NO_3$-N nonpoint source | × | |
| Nitrification (first order) | | × |
| Nitrification (half-saturation) | × | |
| $NH_4$-N assimilation by submerged vegetation | | × |
| $NO_3$-N assimilation by submerged vegetation | | × |
| Denitrification | × | × |
| Reaeration | × | × |
| Photosynthetic oxygen production | × | × |

characteristics a good data fit could be obtained. The main examples of this approach are the CBOD biodegradation rate $K_b$, which depends on the river characteristics (water flow, river bed morphology, depth) as pointed out by Thomann and Mueller (1987), Chapra (1997), Chapra and Pellettier (2003). In the model this parameter has a variable structure, depending on the location, as shown in Table 1. Further, the photosynthetic oxygen production by the submerged vegetation $DO_{ph}(x)$ varies widely, depending on the level of eutrophication. Hence differing levels of oxygen production were introduced after inspecting the river length and assessing the level of vegetative development. The kinetics of the ammonium-N oxidation was also found to vary between a first order and a half-velocity kinetics depending on the river conditions. All these parameter structures, in addition to requiring a careful inspection of the river morphology, introduce a large amount of switching functions into the model, which not only increase its complexity but also make its parameter assessment more difficult, as already point out by Marsili-Libelli et al. (2003) and Checchi et al. (2007).

### 2.3. Point and nonpoint sources

Detecting the effects of nonpoint sources can be a daunting task involving indirect estimation of basin characteristics and land use, as shown by Azzellino et al. (2006). In this study, each point source is modelled by diluting the discharge with the upstream flow and restarting the integration with the new initial conditions. Nonpoint sources of CBOD, $NH_4$-N and photosynthetic oxygen production rate $DO_{ph}(x)$, which are very difficult to detect, were introduced as unknown piecewise constant model inputs and estimated together with proper model parameters. This increases the dimension of the parameter estimation problem and affects the global identifiability by introducing more switching functions. The known point sources along both rivers and the loadings from the wastewater treatment plants are shown in Table 3, whereas their spatial distribution is shown in Fig. 3 for the Sieve and in Fig. 5 for the Ombrone.

## 3. Parameter estimation

The relatively simple model structure of Table 1, together with the step-wise parameter variations, can produce reliable results only if it is supported by a robust estimation procedure. This section is not only concerned with parameter calibration, but also with the identifiability of a variable structure model in which nonpoint sources are included as additional parameters. The modified Simplex algorithm described in Marsili-Libelli (1992) was used as the basic optimization method, together with a more recent theory (Marsili-Libelli et al., 2003) regarding the confidence regions of estimated parameters. These results have been successfully applied to constructed wetlands (Marsili-Libelli and Checchi, 2005) and respirometric models (Checchi and Marsili-Libelli, 2005). Recently a software toolbox named PEAS, an acronym for Parameter Estimation and Accuracy Software, has been published (Checchi et al., 2007). The software is freely available as a Matlab 7 toolkit upon email request to the corresponding author of this paper. It includes all the above theory and implements the complete procedures for parameter estimation and reliability analysis required to calibrate the river quality model of Section 2 and Table 1.

### 3.1. Definition of the parameter estimation problem

Parameter estimation consists of finding the best value of the parameter vector $\boldsymbol{p}$ of $n_p$ dimension in the least-squares

Table 3
Known point sources along the Sieve and Ombrone rivers

| River | Distance (km) | Flow ($m^3\,s^{-1}$) | CBOD ($mg\,L^{-1}$) | DO ($mg\,L^{-1}$) | $NH_4$-N ($mg\,L^{-1}$) | Type of source |
|---|---|---|---|---|---|---|
| Sieve | 0.0 | 3.000 | 5.0 | 12.0 | 0.0 | Upstream conditions |
| | 13.5 | 0.120 | 70.0 | 3.0 | 0.0 | Rabatta WWTP discharge |
| | 18.4 | 0.0125 | 70.0 | 5.0 | 33.0 | Point source discharge at Vicchio |
| | 21.0 | 0.015 | 20.0 | 3.0 | 0.0 | Vicchio WWTP discharge |
| | 30.0 | 0.0126 | 130.0 | 5.0 | 25.0 | Point source discharge at Dicomano |
| | 40.5 | 0.017 | 50.0 | 5.0 | 23.0 | Point source discharge at Rufina |
| Ombrone | 0.0 | 0.3 | 7 | 7 | 1.97 | Upstream conditions |
| | 7.1 | 1.4 | 6.0 | 5.0 | 3.2 | Baciacavallo WWTP discharge |
| | 7.9 | 0.13 | 4.0 | 5.0 | 5.0 | Candeli WWTP discharge |

sense (Seber and Wild, 1989; Marsili-Libelli et al., 2003), i.e. the one which minimizes the weighted quadratic output error functional

$$E(\boldsymbol{p}) = \sum_{j=1}^{q} \sum_{k=1}^{N} \frac{1}{\sigma_{j,k}^2} \left( y_{j,k}^{\exp} - \boldsymbol{y}_{j,k}(\boldsymbol{p}) \right)^2, \tag{5}$$

where $N$ is the number of experimental data for each of the ($j = 1, \ldots, q$) model outputs, $\boldsymbol{y}_{j,k}^{\exp}$ are noisy experimental observations and $\boldsymbol{y}_{j,k}(\boldsymbol{p})$ are the parameter-dependent corresponding model outputs. The index $k$ indicates the locations at which the data were sampled along the river. The weighting factors in Eq. (5) are inversely proportional to the measurement uncertainty $\sigma_{j,k}$ defined in Section 1.2. The optimization of the error functional is solved through a numerical search coupled with the model in Table 1 computing the model output $\boldsymbol{y}$ for each trial of the parameter vector $\boldsymbol{p}$. The Simplex algorithm used in this application is an improved version of the classical flexible polyhedron (Himmelblau, 1972), where the expansion length is optimized in the search direction (Marsili-Libelli, 1992). This optimization method is inherently unconstrained, though in practice model parameters may require constraints, first of all positivity; however, there are several reasons which prevent the introduction of constraints in the Simplex search, as discussed in Checchi and Marsili-Libelli (2005), the most important being that constraints may disrupt its convergence properties and produce a less realistic covariance matrix. On the other hand, if the search is left unconstrained a wrong problem formulation can be easily detected through unrealistic results.

### 3.2. Sensitivity analysis and identifiable parameters

The problem of selecting the best compromise in terms of identifiability, computational complexity and model accuracy is now considered. The need to define a maximum identifiable parameter subset is justified by the data scarcity and the necessity to identify nonpoint sources, which are included as additional parameters and further increases the dimension and the difficulty of the calibration problem.

The selection of identifiable parameter subsets is a particularly relevant subject in environmental system identification (Weijers and Vanrolleghem, 1997; Petersen, 2000; Dochain and Vanrolleghem, 2001; Omlin et al., 2001; Reichert and Vanrolleghem, 2001; Brun et al., 2002; Gernaey et al., 2004; De Pauw, 2005). Almost all these studies are based on trajectory sensitivity functions, defined as the incremental ratio between the variation of each output ($j$) and that of the perturbed parameter ($i$), i.e. $S_{i,j} = \partial y_j / \partial p_i$. The trajectory sensitivities together with the matrices $\boldsymbol{Q}_k = \mathrm{diag}((1/\sigma_{1,k}^2)(1/\sigma_{2,k}^2)\ldots(1/\sigma_{q,k}^2))$ form the Fisher Information Matrix (FIM)

$$\boldsymbol{F} = \sum_{k=1}^{N} S_{i,j}(k)^T \boldsymbol{Q}_k S_{i,j}(k), \tag{6}$$

whose inverse yields the parameter covariance matrix $\boldsymbol{C} = \boldsymbol{F}^{-1}$ (Petersen, 2000; Dochain and Vanrolleghem, 2001; Marsili-Libelli et al., 2003; De Pauw, 2005), thus, maximizing $\boldsymbol{F}$ implies minimizing the estimation error. In fact, by inspecting the trajectory sensitivities the identifiability of the data/model combination can be assessed and optimal experimental conditions can be designed to maximize the estimation accuracy (Insel et al., 2003; De Pauw, 2005; Checchi and Marsili-Libelli, 2005).

The identifiability of a parameter subset is determined according to the following procedure: the parameters are ranked according to a global sensitivity index (Eq. (7)), then a sequence of FIM sub-matrices is formed including an increasing number of parameters starting with the most sensitive; then the OED criteria are tested on each matrix. The largest identifiable subset is then composed of the maximum number of parameters for which the criteria do not show a marked departure from the minimum set ($i = 2$). Most OED criteria, derived from the FIM, were originally conceived as part of a theory of optimal experiments (Fedorov, 1972) with the aim of minimizing the estimates covariance (De Pauw, 2005). Table 4 illustrates the main criteria used in the literature (see e.g. Atkinson and Donev, 1992; Versyck et al., 1998; Petersen, 2000; Insel et al., 2003; De Pauw, 2005; Checchi and Marsili-Libelli, 2005) and their effect on the estimation. For the present study the most effective proved to be **A**, **E** and **mod E**.

To apply these criteria, first the sensitivity matrix $\boldsymbol{S}$ is used to rank the parameters according to their identifiability by defining the global sensitivity index:

$$\xi_i = \sqrt{\frac{1}{N} \sum_{j=1}^{q} \sum_{k=1}^{N} S_{i,j}(k)^2} \quad i = 1, \ldots, n_p, \tag{7}$$

which is similar to that introduced by Reichert and Vanrolleghem (2001) and De Pauw (2005). After reordering the rows of $\boldsymbol{S}$ for a descending $\xi$, a collection of FIM sub-matrices is formed by the principal minors of order $i = 2, \ldots, n_p$, i.e.then the OED criteria are computed for the set $\boldsymbol{F}_i$ ($i = 2, \ldots, n_p$). The maximum identifiable subset includes the first $i$ parameters for which the criteria still yield an acceptable value. This procedure was applied to the Sieve and Ombrone river

Table 4
Optimal experiment design criteria based on FIM

| Criteria | Method | Effect |
|---|---|---|
| **A** | Min (tr $\boldsymbol{F}^{-1}$) | Minimization of the arithmetic mean of parameter errors. |
| **mod A** | Max (tr $\boldsymbol{F}$) | Same as **A**, but insensitive to **FIM** ill-conditioning |
| **D** | Max (det $\boldsymbol{F}$) | Maximizes the volume of the confidence ellipsoid |
| **E** | Max ($\lambda_{\min}$) | Maximizes the minimum eigenvalues of $\boldsymbol{F}$, which is proportional to the length of the largest axis of the confidence ellipsoid |
| **mod E** | Min ($\lambda_{\max}/\lambda_{\min}$) | Minimizes the condition number |

$$\boldsymbol{F} = \begin{bmatrix} F_{1,1} & \dots & F_{1,i} & \dots & F_{1,n_p} \\ \dots & \dots & \dots & \dots & \dots \\ F_{i,1} & \dots & F_{i,i} & \dots & F_{i,n_p} \\ \dots & \dots & \dots & \dots & \dots \\ F_{n_p,1} & \dots & F_{n_p,i} & \dots & F_{n_p,n_p} \end{bmatrix}, \tag{8}$$

models and the ranking of Table 5 was obtained. Applying these criteria to the set $\boldsymbol{F}_i$ computed for the two rivers, the graphs of Figs. 7 and 8 were obtained. In both cases all the criteria indicate that the four most sensitive parameters form an identifiable subset and that the inclusion of the nonpoint sources as additional parameters decreases the overall identifiability. On the other hand, these parameters have to be included in the identification, lacking any prior knowledge about the possible nonpoint sources. Thus the identifiable subset is formed by the "core" parameters plus the nonpoint sources. These subsets for the two rivers are indicated by the shaded rows of Table 5. For the remaining parameters, an adaptation of literature values based on prior knowledge was used, particularly for the reaeration constant $K_r$ (Hornberger and Kelly, 1975; Thomann and Mueller, 1987; Aalderink and Jovin, 1997; Chapra, 1997; Izagirre et al., 2005). The actual parameter estimation and the uncertainty analysis were carried out with the use of the PEAS toolbox (Checchi et al., 2007).

### 3.3. Monte Carlo uncertainty analysis

The statistics of the estimates include at least their confidence intervals, but can be extended to include the full confidence regions (Bates and Watts, 1988; Seber and Wild, 1989) and a previous paper (Marsili-Libelli et al., 2003) illustrates a method for determining the confidence regions of nonlinear models. In this study, however, the application of that theory is prevented by the patched structure of several model parameters. As a robust alternative the Monte Carlo method has been applied to determine the relative uncertainty of the estimate, whose confidence limits are inversely proportional to its identifiability. After a parameter estimate $\hat{\boldsymbol{p}}$ has been determined; the Monte Carlo method generates a parameter distribution based on a large number ($N_{\text{simul}}$) of estimations obtained from perturbed observations $\tilde{y} = y(\hat{\boldsymbol{p}})(1+\varepsilon)$, where $\varepsilon$ is sampled from a Gaussian pdf $N(0, s^2)$ produced

Table 5
Sensitivity scores of the model parameters for the two rivers computed with Eq. (7)

| River | Parameter | $\xi$ | |
|---|---|---|---|
| Sieve | $K_{sa}$ | 791.25 | Core parameters |
| | $K_{a\_max}$ | 376.76 | |
| | $K_b$ | 181.65 | |
| | $K_o$ | 69.79 | |
| | $B_d$ | 49.49 | Nonpoint sources |
| | $NO_{3d}$ | 25.29 | |
| | $DO_{ph}$ | 10.77 | |
| | $K_c$ | 5.28 | Literature or euristic values |
| | $K_{ao}$ | 0.43 | |
| Ombrone | $K_{a\_max}$ | 313.68 | Core parameters |
| | $K_b$ | 198.46 | |
| | $K_{al}$ | 150.98 | |
| | $K_o$ | 44.61 | |
| | $DO_{ph}$ | 42.38 | Nonpoint source |
| | $K_c$ | 17.29 | Literature or euristic values |
| | $K_f$ | 13.68 | |
| | $K_{os}$ | 11.12 | |
| | $K_{ao}$ | 4.90 | |

Only the "core" parameters in the shaded areas were actually calibrated. The other parameter values were either adapted from the literature or obtained from direct sources.

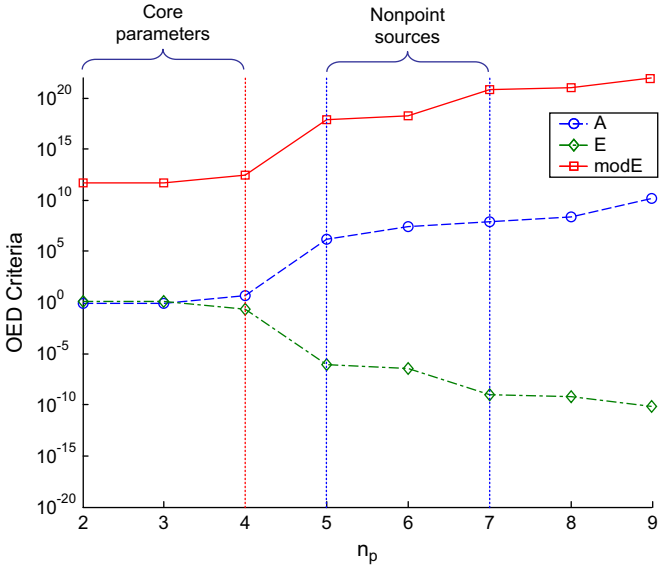


Fig. 7. Identifiability indexes for the Sieve river as a function of the number of estimated parameters $n_p$. The core parameters ($K_{sa}$, $K_{a\_max}$, $K_{bm}$ and $K_o$) are indexed from 1 to 4, whereas indexes from 5 to 7 refer to the nonpoint sources $B_d$, $NO_{3d}$, $DO_{ph}$.

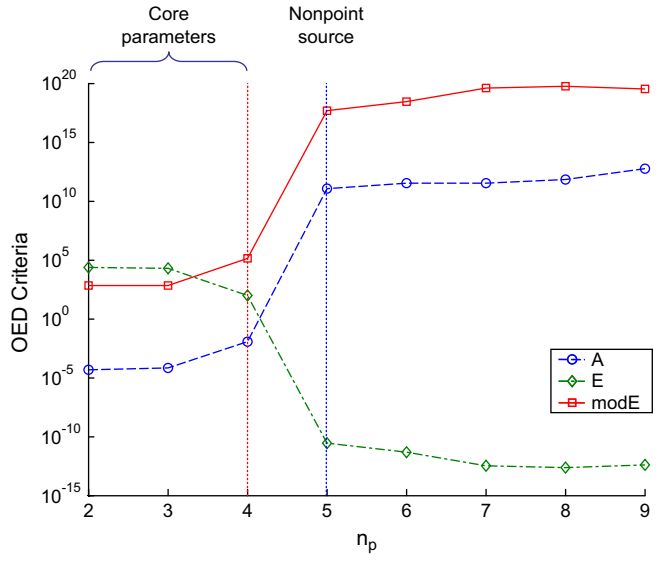


Fig. 8. Identifiability indexes for the Ombrone river as a function of the number of estimated parameters $n_p$. The core parameters ($K_{a\_max}$, $K_b$, $K_{al}$ and $K_o$) are indexed from 1 to 4, whereas index 5 refers to the nonpoint source $DO_{ph}$.

by the built-in Matlab random number generator `randn`. A Gaussian pdf $N(\widehat{m}_i, \widehat{s}_i^2)$ is then fitted to the histogram of the $i$-th estimate as

$$\widehat{p}_i \cong \widehat{m}_i; \delta p_i \cong \pm t^{\alpha/2}_{N_{\text{simul}}-1}\, \widehat{s}_i, \tag{9}$$

where $t^{\alpha/2}_{N_{\text{simul}}-1}$ is the two-tails student's $t$ distribution for the given confidence level $100(1-\alpha)\%$ and $N_{\text{simul}} - 1$ is the degrees of freedom, and $\widehat{s}_i^2 = (1/N_{\text{simul}} - 1)\sum_{j=1}^{N_{\text{simul}}}(p_j - \widehat{m}_i)^2$ is the estimated variance of the Gaussian pdf fitted to the histogram of the $i$-th parameter. If the histogram is mainly composed of narrow central bins, fitting a Gaussian pdf tends to overestimate the variances and hence the confidence limits computed with Eq. (9). An example of the histograms obtained from the Monte Carlo and fitted Gaussian distributions analysis is shown in Fig. 9. For some parameters (e.g. $K_b$) the computed distribution is well approximated with a Gaussian pdf, whereas for others (e.g. $K_{sa}$ and $k_o$) this approximation is overly conservative.

### 3.4. Calibration results

The identified parameters for the two rivers are listed in Table 6 for the Sieve and Table 7 for the Ombrone. The large confidence intervals computed with Eq. (9) confirm the detrimental effect of including the nonpoint sources in the identification. The model responses are shown in Figs. 10 and 11 for the Sieve and Figs. 12 and 13 for the Ombrone. The river behaviour is well explained by the model in both instances, with the point and nonpoint sources playing a crucial role in determining the water quality and the many untreated discharges still representing a major concern. For the Sieve river the temperature gradient is clearly visible in Fig. 10 through the decreasing value of the oxygen saturation concentration. Further, the effect of the nonpoint source downstream of Dicomano is considerable, causing an increase in CBOD in the last 15 km. Fig. 11 shows the nitrate build-up as a consequence of river nitrification, followed by denitrification in the downstream part, favoured by the thick riparian zones.

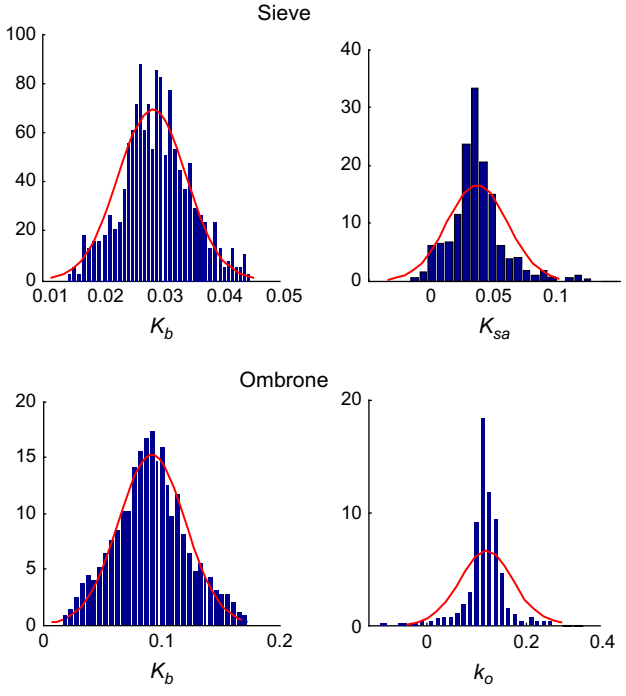


Fig. 9. An example of the distributions of the estimated parameters for the two rivers obtained with Monte Carlo simulations. For some parameters (e.g. $K_b$) the computed distribution is well approximated with a Gaussian function, in other instances (e.g. $K_{sa}$ and $k_o$) the variance of the approximating Gaussian distribution is considerably larger than that of the original histogram, causing an overestimation of the corresponding confidence interval.

Table 6
Calibrated parameters of the Sieve river model

| Parameter | Estimated value and Monte Carlo confidence limits | Length (km) |
|---|---|---|
| $K_{sa}$ | $0.037 \pm 0.032$ | $0 < x < 48.4$ |
| $K_{a\_max}$ | $0.048 \pm 0.089$ | $0 < x < 48.4$ |
| $K_{bm}$ | $0.029 \pm 0.008$ | $0 < x < 48.4$ |
| $K_o$ | $0.139 \pm 0.073$ | $0 < x < 48.4$ |
| $B_d$ | $0.321 \pm 0.282$ | $35 < x < 48.4$ |
| $NO_{3d}$ | $0.385 \pm 0.158$ | $8.2 < x < 21.4$ |
| $DO_{ph}$ | $2.11 \pm 0.281$ | $0 < x < 10$ |
| | Constant value | Reference |
| $Y_a$ | 0.24 | Brown and Barnwell (1987), Chapra (1997) |
| $K_c$ | 1.5 | Chapra (1997), Thomann and Mueller (1987) |
| $K_{aa}$ | 0.075 | Chapra (1997) |

Table 7
Calibrated parameters of the Ombrone river model

| Parameter | Estimated value and Monte Carlo confidence limits | Length (km) |
|---|---|---|
| $K_b$ | $0.109 \pm 0.043$ | $0 < x < 10.5$ |
| | $0.009 \pm 0.001$ | $10.5 < x < 14.6$ |
| $K_{al}$ | $0.512 \pm 0.106$ | $0 < x < 10.5$ |
| | $0.093 \pm 0.092$ | $10.5 < x < 14.6$ |
| $K_{a\_max}$ | $0.117 \pm 0.031$ | $0 < x < 10.5$ |
| | $0.137 \pm 0.064$ | $10.5 < x < 14.6$ |
| $K_o$ | $0.117 \pm 0.077$ | $0 < x < 10.5$ |
| | $0.137 \pm 0.093$ | $10.5 < x < 14.6$ |
| $DO_{ph}$ | $1.656 \pm 0.279$ | $0 < x < 7.1$ |
| | $1.760 \pm 0.684$ | $7.1 < x < 10.4$ |
| | $3.847 \pm 0.632$ | $10.4 < x < 14.6$ |
| | Constant value | Reference |
| $Y_a$ | 0.24 | Brown and Barnwell (1987), Chapra (1997) |
| $K_f$ | 0.026 | Lindenschmidt (2006) |
| $K_c$ | 0.85 $0 < x < 10.5$ | Chapra (1997), Thomann and Mueller (1987) |
| | 0.67 $10.5 < x < 14.6$ | |
| $\delta$ | 0.2 | Brown and Barnwell (1987) |
| $C_t$ | 0.1 | Jørgensen and Bendoricchio (2001) |
| $pK_1$ | 5 | Jørgensen and Bendoricchio (2001) |
| $pK_2$ | 8.8 | Jørgensen and Bendoricchio (2001) |
| $K_{os}$ | 1.34 | Lindenschmidt (2006) |

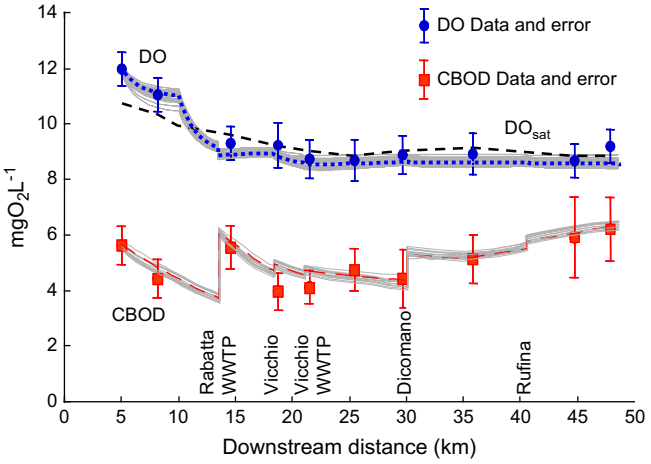


Fig. 10. Calibrated model response for the Sieve river: CBOD/DO dynamics. The thin grey lines represent model responses with altered parameters.

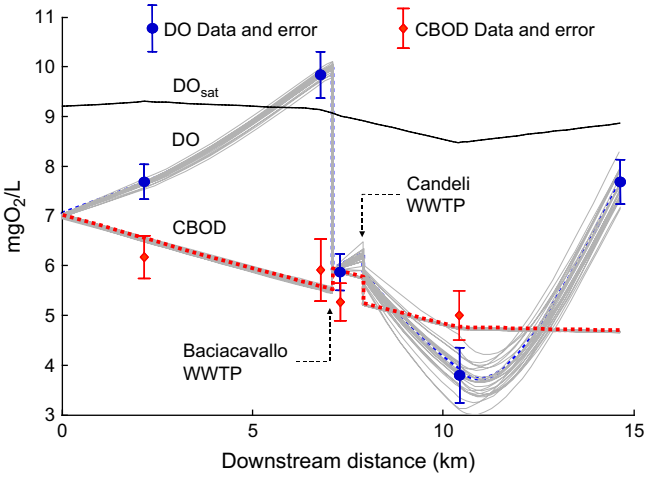


Fig. 12. Calibrated model response for the Ombrone river: BOD/DO dynamics. The thin grey lines represent model responses with altered parameters. The error margin for the CBOD data is larger than for the Sieve due to less campaign and larger inherent data variability. The abrupt DO decrease at 10 km is due to the mixing of the deoxygenated Baciacavallo WWTP discharge.

For the Ombrone river, Fig. 12 shows the DO abrupt decrease caused by the oxygen-depleted massive inflow from the Baciacavallo WWTP, though the corresponding CBOD increase is limited. The DO sag downstream of the discharge point is also clearly visible. Again, the insufficient nitrogen removal is visible in Fig. 13 with a sharp increase at the discharge point, followed by nitrification and algal uptake in the downstream reach. The same figure shows the continuing denitrification and nitrate uptake of the high upstream level, partly offset by the Baciacavallo discharge.

Though these results show that the model response always falls within the experimental error boundaries, it should be considered that given the small number of data points and the large number of parameters it is not difficult to obtain a good average fit. In order to assess the model validity further, a multiple parameter sensitivity analysis, as described by Haefner (2005), was performed. A number of simulations were produced with parameters drawn from independent Gaussian distributions with mean equal to the calibrated value and variance given by the squared confidence boundaries obtained by the Monte Carlo analysis of Section 3.3. These results are shown as clusters of thin grey lines in the model response figures (Figs. 10 and 11 for the Sieve and Figs. 12 and 13 for the Ombrone). It can be seen that the envelope of these trajectories is almost always within the experimental boundaries. The only moderate exceptions are the nitrate response in the Sieve river and the second DO sag in the Ombrone river. The first instance can be explained with the actual variability of the nonpoint $NO_3$ input, whereas in the Ombrone case, the photosynthetic oxygen production in that reach was found to represent the largest contribution to model

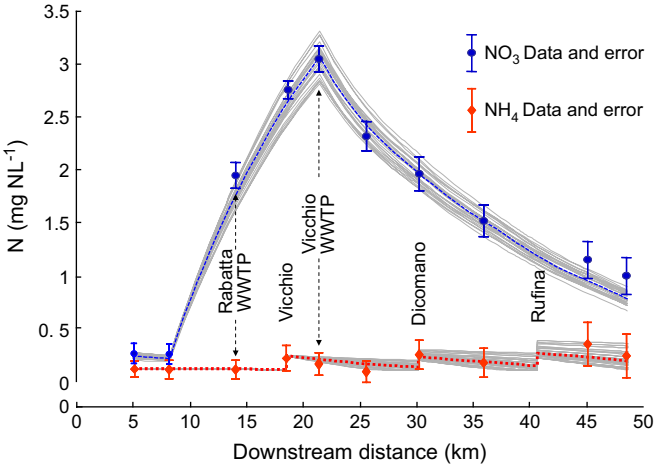


Fig. 11. Calibrated model response for the Sieve river: nitrogen dynamics. The thin grey lines represent model responses with altered parameters.

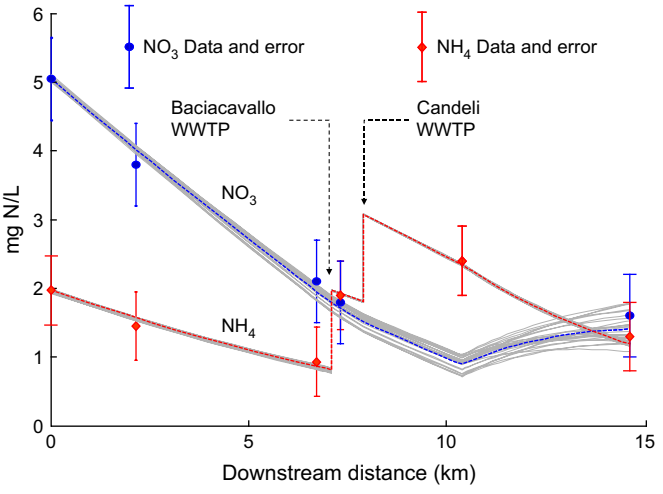


Fig. 13. Calibrated model response for the Ombrone river: nitrogen dynamics. The thin grey lines represent model responses with altered parameters. The effect of the two WWTPs is an abrupt increase in $NH_4$-N, which is then metabolized with a faster rate than in the upstream part due to increased algal population. The reduced $NO_3$-N conversion in the downstream part is due to an increased algal $NH_4$-N uptake.

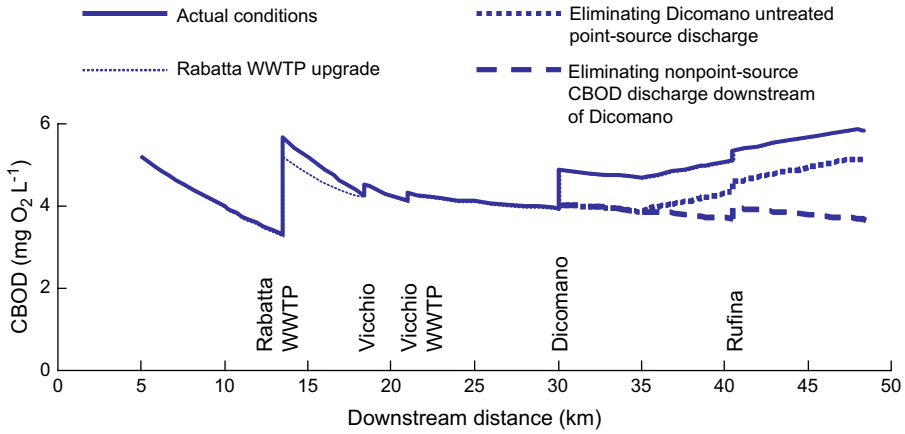


Fig. 14. Effect of all the possible upgrading options in the Sieve river: improving the Rabatta WWTP has only a limited effect, whereas eliminating either the Dicomano untreated point source discharge or the downstream nonpoint source of CBOD has a considerable impact on the downstream water quality, with the latter having the largest effect.

uncertainty. A similar approach was used by Vandenberghe et al. (2007) who complemented the average model response with the uncertainty range (95 and 5% percentiles) due to parameter uncertainty.

## 4. Scenario generation

The calibrated model can be used to assess the effectiveness of management decisions regarding the unsolved problems affecting the basin: upgrade the existing wastewater treatment plants and/or control the nonpoint source loading. The effect of these actions is shown in Fig. 14, halving the output CBOD concentration of the Rabatta WWTP with respect to the values of Table 3 or removing either the untreated discharge at Dicomano or the nonpoint source CBOD downstream of that settlement. It can be seen from Fig. 14 that the effect of the Rabatta WWTP discharge is quite limited in space, whereas eliminating either the untreated source in Dicomano or the nonpoint source downstream of Rufina has a major impact on the entire final reach of the Sieve, with the latter action producing the larger improvement.

## 5. Conclusion

The goal of this study was to set up a simple model to describe the stationary water quality in small river basins where the use of full-fledged models, such as those distributed by USEPA, is too complex from a practical and economic viewpoint. After introducing the basic model structure, it was shown how this could be adapted to rivers of widely differing characteristics. Given the short river length compared to the extent required by the self-purification dynamics, point and nonpoint sources play a key role in shaping the model response and have to be accounted for, either by direct inspection or parametric estimation. Further, the varying river characteristics, in terms of morphology, hydraulics and vegetation, require the introduction of variable parameters, thus complicating the originally simple model structure. To determine the identifiability of the resulting model an identifiability assessment was carried out based on sensitivity analysis and OED criteria. The identifiable subset was determined by ranking the parameters in terms of decreasing sensitivity and computing the associated Fisher Information Matrices. Having determined an identifiable ‘‘core’’ parameter subset, it was found that the inclusion of the nonpoint sources as additional parameters affected the identifiability to a considerable extent. However, the combined parameter–sources calibration was made possible by the use of a robust estimation algorithm, which also provided an estimate of the confidence limits through Monte Carlo analysis. The calibrated model responses and their sensitivity envelopes are in good agreement with the observed water quality data, therefore the model is consistent and can be used to generate scenarios as a part of a general strategy to conserve or improve the water quality.

## Acknowledgement

This research was partly supported by the Regional Agency for Environmental Protection (ARPAT) through cooperation with the National Research Centre for Inland and Coastal Waters (CTN-AIM). The authors also wish to acknowledge the cooperation of the ARPAT provincial laboratory of Prato for their support in the Ombrone river sampling.

## References

Aalderink, R.H., Jovin, J., 1997. Identification of the parameters describing primary production from continuous oxygen signals. Water Science and Technology 36, 43–51.

Atkinson, A.C., Donev, A.N., 1992. Optimum Experimental Designs. Clarendon Press, Oxford, 352 pp.

Azzellino, A., Salvetti, R., Vismara, R., Bonomo, L., 2006. Combined use of the EPA-QUAL2E simulation model and factor analysis to assess the source apportionment of point and nonpoint loads of nutrients to surface waters. Science of the Total Environment 371, 214–222.

Bates, D., Watts, D., 1988. Nonlinear Regression Analysis and Its Applications. John Wiley & Sons, New York, 365 pp.

Beck, M.B., 1987. Water quality modelling: a review of the analysis of uncertainty. Water Resources Research 23, 1393–1442.

Brown, L.C., Barnwell, T.O., 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS. EPA/600/3-87-007. U.S. Environmental Protection Agency, Athens, GA, 189 pp.

Brun, R., Kuhni, M., Siegrist, H., Gujer, W., Reichert, P., 2002. Practical identifiability of ASM2d parameters – systematic selection and tuning of parameter subsets. Water Research 36, 4113–4127.

Chapra, S., Pellettier, G., 2003. QUAL2K: A Modeling Framework for Simulating River and Stream Water Quality. Documentation and Users Manual. Civil and Environmental Engineering Dept., Tufts University, Medford, MA. Available from: http://www.epa.gov/athens/wwqtsc/html/qual2k.html

Chapra, S., 1997. Surface Water-quality Modeling. McGraw-Hill Int. Ed., New York, 844 pp.

Checchi, N., Marsili-Libelli, S., 2005. Reliability of parameter estimation in respirometric models. Water Research 39, 3686–3696.

Checchi, N., Giusti, E., Marsili-Libelli, S., 2007. PEAS: a toolbox to assess the accuracy of estimated parameters in environmental models. Environmental Modelling and Software 22, 899–913.

De Pauw, D.J.W., 2005. Optimal Experimental Design for Calibration of Bioprocess Models: A Validated Software Toolbox. PhD thesis in Applied Biological Sciences, BIOMATH, University of Gent. Available from: <http://biomath.ugent.be/publications/download/>.

Dochain, D., Vanrolleghem, P.A., 2001. Dynamical Modelling and Estimation in Wastewater Treatment Processes. IWA Publishing, London, 342 pp.

Fedorov, V.V., 1972. Theory of Optimal Experiments. Academic Press, New York, 292 pp.

Gernaey, K.V., van Loosdrecht, M.C.M., Henze, M., Lind, M., Jørgensen, S.B., 2004. Activated sludge wastewater treatment plant modelling and simulation: state of the art. Environmental Modelling and Software 19, 763–783.

Haefner, J.W., 2005. Modeling Biological Systems, Principles and Applications. Springer, New York, 475 pp.

Halfon, E., 1983. Is there a best model structure? 1. Modelling the fate of a toxic substance in a lake. Ecological Modelling 20, 135–152.

Himmelblau, D., 1972. Applied Nonlinear Programming. McGraw-Hill, New York, 498 pp.

Hornberger, G.M., Kelly, M.G., 1975. Atmospheric reaeration in a river using productivity analysis. ASCE Journal of Environmental Engineering 101, 729–739.

Insel, G., Orhon, D., Vanrolleghem, P.A., 2003. Identification and modelling of aerobic hydrolysis mechanisms – application of optimal experimental design. Journal of Chemical Technology and Biotechnology 78, 437–445.

Izagirre, O., Bermejo, M., Pozo, J., Elosegi, A., 2005. RIVERMET: an excel-based tool to calculate river metabolism from diel oxygen-concentration curves. Environmental Modelling and Software 22, 24–32.

Jørgensen, S.E., Bendoricchio, G., 2001. Fundamentals of Ecological Modelling. Elsevier, Amsterdam, 530 pp.

Lindenschmidt, K.E., 2006. The effect of complexity on parameter sensitivity and model uncertainty in river water quality modelling. Ecological Modelling 190, 72–86.

Marsili-Libelli, S., 1992. Parameter estimation of ecological models. Ecological Modelling 62, 233–258.

Marsili-Libelli, S., Checchi, N., 2005. Identification of dynamic models for horizontal subsurface constructed wetlands. Ecological Modelling 187, 201–218.

Marsili-Libelli, S., Guerrizio, S., Checchi, N., 2003. Confidence regions of estimated parameters for ecological systems. Ecological Modelling 165, 127–146.

Omlin, M., Brun, R., Reichert, P., 2001. Biogeochemical model of lake Zurich: sensitivity, identifiability and uncertainty analysis. Ecological Modelling 141, 105–123.

Pearl, J., 1978. On the connection between the complexity and credibility of inferred models. International Journal of General Systems 4, 255–264.

Petersen, B., 2000. Calibration, Identifiability and Optimal Experimental Design of Activated Sludge Models. PhD thesis, Ghent University, Ghent, Belgium. Available from: <http://biomath.ugent.be/publications/download/>.

Reichert, P., Vanrolleghem, P.A., 2001. Identifiability and uncertainty analysis of the River Water Quality Model No. 1 (RWQM1). Water Science and Technology 43 (7), 329–338.

Reichert, P., Borchardt, D., Henze, M., Rauch, W., Shanahan, P., Somlyódy, L., Vanrolleghem, P.A., 2001. River Water Quality Model No. 1. Scientific and Technical Report No. 12. IWA Publishing, London, 131 pp.

Rinaldi, S., Soncini-Sessa, R., Stehfest, H., Tamura, H., 1979. Modelling and Control of River Quality. McGraw-Hill, New York, 380 pp.

Seber, G.A., Wild, C.J., 1989. Nonlinear Regression. John Wiley & Sons, New York, 768 pp.

Snowling, S.D., Kramer, J.R., 2001. Evaluating modelling uncertainty for model selection. Ecological Modelling 138, 17–30.

Thomann, R.V., Mueller, J.A., 1987. Principles of Surface Water Quality Modeling and Control. Harper & Row, New York, 644 pp.

Vandenberghe, V., Bauwens, W., Vanrolleghem, P.A., 2007. Evaluation of uncertainty propagation into river water quality predictions to guide future monitoring campaigns. Environmental Modelling and Software 22, 725–732.

Versyck, K.J., Claes, J.E., van Impe, J.F., 1998. Optimal experimental design for practical identification of unstructured growth models. Mathematics and Computers in Simulation 46, 621–629.

Weijers, S.R., Vanrolleghem, P.A., 1997. A procedure for selecting best identifiable parameters in calibrating activated sludge model no.1 to full-scale plant data. Water Science and Technology 36 (5), 69–79.

Wool, T.A., Ambrose, R.B., Martin, J.L., Comer, E.A., 2006. Water Quality Analysis Simulation Program (WASP) Version 6.0 DRAFT: User's Manual. U.S. Environmental Protection Agency, Athens, GA. Available from: <http://www.epa.gov/athens/wwqtsc/html/wasp.html>.