Water Pollution Models based on Stochastic Differential Equations

Zhiyong HUANG, Hiroshi MORIMOTO

Department of Earth and Environmental Sciences
Graduate School of Environmental Studies
Nagoya University

Abstract

Water pollution has been a crucial problem in many countries and has attracted researcher's attention from all over the world.

In a non-tidal river, from the beginning of material input, mixing and reaction, to the end of material output, the system considered should be dynamic and stochastic.

The stochastic differential equation (SDE) is a theory that studies stochastic process. It can describe the physical relationships of different factors, and can predict the effects of future loading and control policies on the environment. It has the advantage of water pollution control. This would then allow an attempt to formulate a general model on self-purification non-tidal rivers.

Most water pollution models have a structure of input/output system. However these models do not assume explicit knowledge of pre-biochemical reaction, such as the explicit pre-reaction concentration. An internally descriptive model exploits the available information on the phenomena determining the system's behaviours, e.g. the physical and biochemical mechanisms which control the internal descriptions. To exploit the information of pre-biochemical reactions is an essential task in modeling.

This paper firstly, designs a structure of input/mixing/output model (we will call it the three-steps model); secondly, it represents each process of material input, mixing and output by an SDE respectively, where the additional equations representing mixing processes exploit the information of pre-biochemical reactions; and thirdly, it shows that the three-steps model has the advantage of prediction and control, using the numerical solutions of SDEs of the three-step model.

The stochastic numerical methods applied here are discrete time approximation methods subject to the Ito-Taylor expansion schemes.

1 Introduction

Our environment is worsening and our water supply is decreasing, however our dependence on this natural resource remains. Every year, we need 270,100 million tons more water because of 90 million increasing population. Every year, approximately 25 million persons die as a result of water pollution. To protect the environment and keep water free of pollution is a global responsibility.

Although various wastewater treatment plants are established deal with water pollution, most of the wastewater is discharged directly into rivers or lakes. Therefore, it is important for us to understand what the mechanism in a natural purification system is. To understand how to control and predict water quality becomes a project of our studying. Having considered the objectives, a model must be developed to represent the system.

A non-tidal self-purification river system is a basic system for studying the self-purification system.

When assessing the quality of water in a non-tidal self-purification river, many factors should be considered: the level of dissolved oxygen (DO); the presence of nitrates, chlorides, phosphates; the level of suspended solids; environmental hormone; chemical oxygen demand (COD), such as heavy metals, and the presence of bacteria. However, in a study about general water quality, it would be difficult to consider in detail all of these factors. Therefore we should choose some principal measures of water quality that will indicate the general quality of the river. In this paper, four factors are considered: DO, biochemical oxygen demand (BOD), substrate, and sludge flow rate.

A good water pollution model on a non-tidal river should possess the following properties.

(1) It should be a dynamic (time varying) or stochastic model. From the beginning of the material input, mixing and reaction, to the end of material output, the system is dynamic and stochastic on inner components, such as turbulent influx in feed-rates, random perturbations in concentrations; and the system changes with changing factors, such as wind gust, varying temperatures and sunlight.

(2) It should adequately characterize the important stochastic, dynamic and steady state aspects of the system behaviour, such as a proper structure of model for representing the system's process

(3) If possible, it should provide a reasonable mathematical approximation of the physical biochemical changes (interaction) occurring in the river system. This means the degree and structure of interaction between variables and/or model components should be identified [2].

Research on the modeling of the BOD-DO interaction in a non-tidal river has been dominated by the classical model of Streeter and Phelps (1925), which first appeared in 1925, and has been improved in various ways by other investigators such as Dobbins and

Camp, Peter Young and Bruce Beck, C.J.Harris, etc. [5].

The class of continuous-time, internally descriptive models are all derived from the application of mass balanced principles to the river stretch. The analysis of these models is well suited in the background of chemical process theory (see, e.g. Himmelblau and Bischoff, (1968)), since, after all, the behaviour of a reach of river can be viewed along lines similar to that of a chemical reactor.

By the criterion of the quantity of the fluid mixing, the physical bio-chemical models are distinguish as follows [1][5]:

(1) No mixing (Zero diffusion)

$$\frac{\partial c}{\partial t} = f_1(\frac{\partial c}{\partial x}, c(x,t)).$$

(2) Partial mixing (turbulent type diffusion)

$$\frac{\partial c}{\partial t} = f_2(\frac{\partial^2 c}{\partial x^2}, \frac{\partial c}{\partial x}, c(x,t)).$$

(3) Intimate mixing, with simple transport delay time T_d between spatial points x_0, x_1. (Peter Young and Bruce Beck (1974)).

$$\frac{\partial c}{\partial t} = f_3(c(x_1,t), c(x_0, t - T_d)).$$

In chemical engineering terminology the model type (3) corresponds to continuously stirred tank reactors (CSTR, Himmelblau and Bischoff, (1968)) plus transportation delay idealization of a reach river.

(4) SDEs type;

$$dc_t = a(c,t)dt + b(c,t)dW_t.$$

with their appropriate boundary conditions.

Where $W_{(t)}$ is a Wiener process. In detail, see section 2.

Functions $f_i, i = 1, 2, 3.$ are in general, nonlinear functions of the substance concentration c.

All of these models are dynamic, applying mass balance law.

From the theoretical point of view:

a) Model type (2) would appear to represent the system dynamics most accurately. It counts the values of material concentration reacting both from the time state and from the variable locations in space along the river stretch.

b) Model type (3) may appear to be a strange idealization of the river system. It counts the values of material concentration reaction simplifying by minus a delay time T_d.

c) In model type (4) e.g. Harris' model, by a noise term, it lumps the space parameters and accounts for the system errors and measurement errors. It just simplifies the representation. For controllable, predictive and practical dynamic investigations the lumped parameter model type (4) seems the most attractive, because it takes advantage of the SDEs theory which is a theory studying stochastic process. It can describe the physical relationships of different factors, and can predict the effects of future loading and controlling policies on the environment.

SDEs will be of great benefit to water pollution control and study.

Although these models have a structure of input/output, they do not assume explicit knowledge of pre-biochemical reactions, such as the explicit pre-reaction concentration.

In this paper, we try to design a structure of the input/mixing/output model and call it the three-steps model. These three steps can be represented respectively by three stochastic differential equations (SDEs). For these three steps we can take random input as a dynamic process subject to turbulent input only; and take substance-mixing process as a continuously stirred process so that it contributes to construct a relatively stable state for biochemical reaction processes, and by a mixing process, it produces an average value of concentration for calculating the result of biochemical reactions.

2 Harris' input/output model

The Harris' model described by SDEs (model type (4)) is used as the starting point for this study [1].

A banded DO-BOD model:

$$\begin{aligned} dq^i/V^i &= (\alpha^i + 1 - \beta^i \cdot q^i/V^i)dt + (2\gamma^i \cdot q^i/V^i)^{1/2} \cdot dW_t \quad (input) \\ dc_d^i &= q^i/V^i\{(c_d^{i-1} - c_d^i) + K_2^i \cdot (c_{ds}^i - c_d^i) - K_1^i \cdot C_b^i + m^i + \delta_1^i\} \cdot dt \quad (DO) \\ dc_b^i &= q^i/V^i\{(c_b^{i-1} - c_b^i) - K_1^i \cdot (c_b^{i-1} - c_b^i) - K_2^i \cdot C_d^i + R^i + \delta_2^i\} \cdot dt \quad (BOD) \end{aligned}$$

where

i: the scalar of i-th reach of a river,

$W_(t)$: a Wiener process,

q^i: sludge volume flow rate;

V^i: volume of i-th reach of river,

c_b^i: BOD concentration,

c_d^i : DO concentration,

c_{ds}^i: saturation concentration of DO,

K_2^i: DO re-aeration rate,

K_1^i: BOD decay rate,

m^i: net rate of removal of DO by mud deposits,

R^i: the addition of BOD due to local runoff from adjoining land and effuent discharge from sewage plants,

δ_1^i and δ_2^i,: accounted for photosynthetic effects on algae populations,

q^i: evaluated so that it always remains non-negative (for physical compatibility) and is subjected to band limited perturbations (filtered white noise). It can be demonstrated for the above models that q^i is stationary Markov process with decay parameter β^i, mean value $(\alpha^i + 1)/\beta^i$ and variance $\gamma^i(\alpha^i + 1)/\beta^i$. These parameters allow adequate modeling for a particular physical environmental configuration, and may be found by standard statistical estimation techniques.

Equations of BOD and DO are ordinary differential equations but are modeled as SDEs in the above, due to their functional relationships with the random component q^i.

If Harris' model is applied to a self-purification non-tidal river system, there are two improprieties.

First, during a short time interval dt, the value of input BOD/ DO concentrations appears to improperly act as the value of biochemical reacting. In fact, the average concentration over the time interval dt and over space volume V^i would be a suitable one. The value of input concentration would be above this average level in general.

The reasons are: first, after the material input, the BOD concentration is reduced by a mixing process and a biochemical reaction; second, in the majority of the cases, the parameters used are obtained from laboratory analyses, namely, they are obtained from a quite stable or steady state. Here, the "mean of stable" is similar to "the sense of average".

Usually, there are two kinds of parameter sets, one is obtained from laboratory analyses,(e.g. for k_2^i ,the re-aeration rate constant and k_1^i, the BOD decay rate constant, see Henry's law, or for k_1^i see the discussion of Hansen and Frankel, (1965). (BOD estimation takes 5 days)); another one is obtained from applying parameter estimation techniques to the set of field observations data sets, or those obtained from empirical relationships (e.g. for k_2^i, the re-aeration rate constant, see Owens and colleagues,(1964)) [5].

It would be more suitable to apply the value of average concentration, instead of inputting one to calculate the reaction concentration would be more suitable.

The reasons why we apply the parameters obtained from laboratory analyses are: first, a quite stable or steady state is a basic requirement for modeling and developing; second, we will discuss the model from a theoretical sense rather than verifying it against field data; third, we are interested in what the results will be happened to the two different structures of models if the same parameters and similar equations are applied; and fourth, our goal is to formulate a general model on the self-purification river system.

Second, the influent volume rate of sludge q^i equals the effluent volume rate in Harris' model. If we consider a long stretch of river, it would be inappropriate, because in general, the input volume rate of sludge (wastewater) is not equal to the output one.

How can we solve these two problems?

3 An idea of the wastewater treatment plant

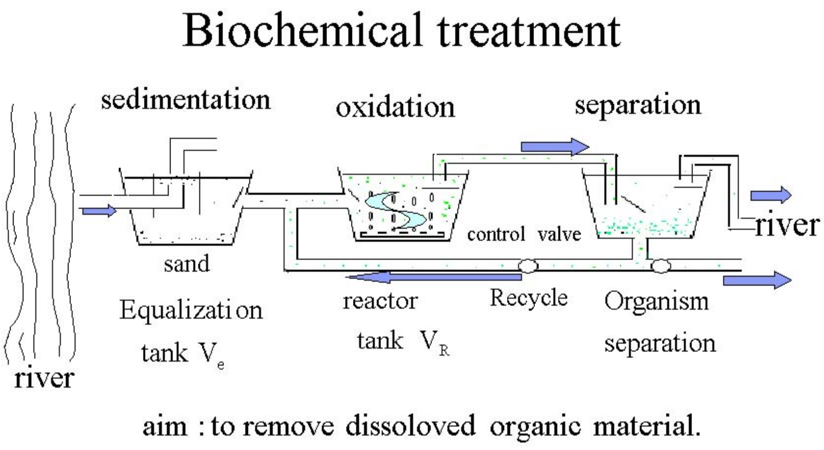

Figure 1: Flow scheme of active sludge control

There are variations of wastewater treatments, in which the activated sludge reactor (since its introduction in 1914), i.e. biochemical treatment seems to be the most active one. Usually, the wastewater treatment plant consists of three tanks, a sedimentation (equalization) tank, a reactor tank and organism separation tank, and a return sludge system as illustrated in figure 1.

The main objective of the sedimentation task is that it removes the most obvious identifiable solids such as grease and floating matter, and reduces BOD and SS (suspended solid) concentrations.

The reactor tank is normally achieved by some from of aerobic biological process, this means the organic matter is degraded by aerobic micro-organisms, and is aimed at further removing the non-settleable, largely organic matter to produce an effluent suitable for discharge into inland streams. To satisfy the demand exerted during reaction process, an oxygen replenishment re-aeration system must be present in sufficient quantities to prevent DO levels from falling to the point where the water cannot support the range of aquatic life.

Organism separation tank is used to remove or re-use dissolved organics, so that the effluent may be rendered suitable for discharge [9].

So far, we know the function of the sedimentation tank has made the system process stable, from a theoretical view.

Based on this knowledge, we can understand that the sedimentation tank, prior to reaction, is used to contribute to process stability in modeling. From this we know that the mixing process is completely prior to the reaction process, i.e. the pre-reaction concentration in the reaction tank equals the mixing concentration in the sedimentation tank.

4 On the three-steps model

Now, we apply the idea that the function of the sedimentation (equalization) tank is a process of fluid mixing, to a self-purification river system. Then we design a new structure of the model as Figure 1, from which, we can respectively study the mixing process and biochemical reaction process closely.

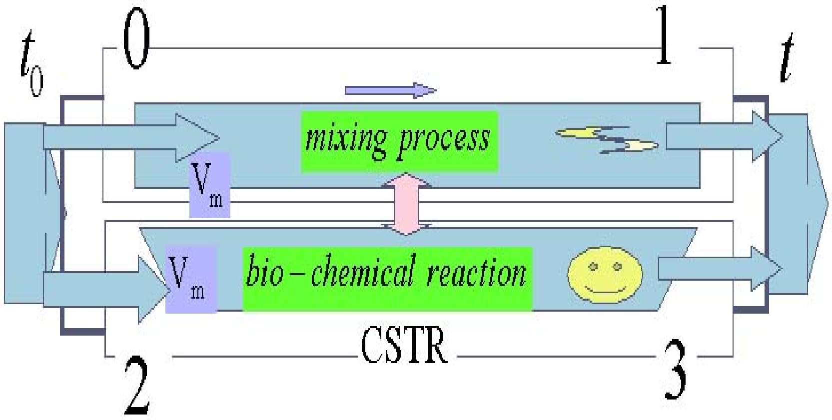

Figure 2: the three-steps model

Figure 2 illustrates the process of material input/mixing/output. (It is called the three-steps model.)

In the three-steps model, physical process (mixing) and biochemical process are united, (i.e. they happen interactively during the same time interval), but they are not stringed together (i.e. they happen one by one at two successive but different time intervals.) It is not like a water treatment plant where the mixing process and biochemical process are stringed together.

More strictly speaking, "united" means that the two processes are interactive and are performed in the same time interval dt. Namely, there is a circle of exchanged process subjects to the value of substance concentration transporting: output value of mixing c_1 $\rightarrow$ input reaction value c_2 $\rightarrow$ output reaction value c_3 $\rightarrow$ a part of the next mixing value $(c_3 \cdot V + c_0 \cdot q_0)/V \rightarrow c_1$ again.

These equations represent the three-steps model, and correspond to the sludge flow rate, the concentration of substrate (pollutant), BOD and DO as follows:

a) The equations of the sludge flow rate and the concentration of substrate (pollutant):

$$dq_0 = (\alpha_1 + 1 - \beta_1 \cdot q_0) \cdot dt + (2\gamma_1 \cdot q_0)^{1/2} \cdot dW_1. \tag{1}$$
$$dV = (q_0 - q_1) \cdot dt. \tag{2}$$
$$dc_{s0} = (\alpha_3 + 1 - \beta_3 \cdot c_{s0})/V \cdot dt + (2\gamma_3 \cdot c_{s0}/V)^{1/2} \cdot dW_3. \tag{3}$$
$$dc_{s1} = q_0 \cdot (c_{s0} - c_{s1})/V \cdot dt. \tag{4}$$
$$V \cdot dc_{s3} = \{q_2 \cdot (c_{s2} - c_{s3}) - K_{51} \cdot V \cdot [K_{b2} \cdot q_2 + K_{53} \cdot (c_{s2} - c_{s3})]\} \cdot dt. \tag{5}$$

where $q_0 = q_2, \quad q_1 = q_3 = E(q_0) = (\alpha_1 + 1)/\beta_1$,
$K_{53} = \Delta c_b/\Delta c_s$=yields constant,
q_0: influent sludge flow rate,
q_2: effluent sludge flow rate,
V: the volume of the river stretch,
c_{s0}: incoming substrate (pollutant) concentration,
K_{b2}: micro-organism concentration,
K_{51}: growth rate constant for micro-organisms,

b) BOD and DO equations:

$$dc_{BOD0} = (\alpha_6 + 1 - \beta_6 \cdot c_{BOD0})/V \cdot dt + (2\gamma_6 \cdot c_{BOD0}/V)^{1/2} \cdot dW_6. \tag{6}$$
$$dc_{BOD1} = q_0/V \cdot (c_{BOD0} - c_{BOD1}) \cdot dt. \tag{7}$$
$$dc_{BOD3} = \{(q_2/V - K_{81}) \cdot (c_{BOD2} - c_{BOD3}) - K_{82} \cdot c_{DO3} + \delta_8\} \cdot dt. \tag{8}$$
$$dT = (\alpha_9 + 1 - \beta_9 \cdot T) \cdot dt + (2\gamma_9 \cdot T)^{1/2} \cdot dW_9. \tag{9}$$
$$c_{DO0} = 482.5/(32.6 + T). \tag{10}$$
$$dc_{DO1} = q_0/V \cdot (c_{DO0} - c_{DO1}) \cdot dt. \tag{11}$$
$$dc_{DO3} = \{q_2/V(c_{DO2} - c_{DO3}) - K_{82}(c_{satu} - c_{DO3}) - K_{81}c_{BOD3} + m_{removal} + \delta_{11}\}dt. \tag{12}$$

where:
T: temperature,
c_{DO}: DO concentration,
c_{DO0}: for benefit of simulation, to show that it is dynamic, where it is assumed to be affected by the factor of temperature only and is evaluated by Henry's law,
c_{BOD}: BOD concentration,
c_{satu}: saturation concentration of DO,
K_{82}: DO re-aeration rate$=0.3 \cdot 1.02^{T-20}$, by Henry's law,
K_{81}: BOD decay rate $=0.17 \cdot 1.047^{T-20}$, by Henry's law,
$m_{removal}$: net rate of removal of DO by mud deposits,

δ_8 and δ_{11}: accounted for photosynthetic effects on algae populations, and set $q_1 = q_0, \quad q_3 = q_2, \quad E(q_0) = (\alpha_1 + 1)/\beta_1.$

Comparing to Harris' model, the term of R which is the addition of BOD due to local runoff from adjoining land effluent discharge from sewage plant. Here it is accounted in the influent equation.

Define the state variables, and rewrite the equations as:
$q_0 = x_1, \quad V = x_2, \quad c_{s0} = x_3, \quad c_{s1} = x_4,$
$c_{s3} = x_5, \quad c_{BOD0} = x_6, \quad c_{BDO1} = x_7, \quad c_{BDO3} = x_8,$
$T = x_0, \quad c_{DO0} = x_9, \quad c_{DO1} = x_{10}, \quad c_{DO3} = x_{11},$

$$dx_1 = (\alpha_1 + 1 - \beta_1 \cdot x_1)dt + (2\gamma_1 \cdot x_1)^{1/2} dW_1. \tag{13}$$

$$dx_2 = (x_1 - q_1)dt. \tag{14}$$

$$dx_3 = (\alpha_3 + 1 - \beta_3 \cdot x_3)/x_2 dt + (2\gamma_3 \cdot x_3/x_1)^{1/2} dW_3. \tag{15}$$

$$dx_4 = x_1 \cdot (x_3 - x_4)/x_2 dt. \tag{16}$$

$$dx_5 = \{q_2 \cdot (x_4 - x_5) - K_{51} \cdot x_2 \cdot [K_{b2} \cdot q_2 + K_{53} \cdot (x_4 - x_5)]\}/x_2 dt. \tag{17}$$

$$dx_6 = (\alpha_6 + 1 - \beta_6 \cdot x_6)/x_2 dt + (2\gamma_6 \cdot x_6/x_2)^{1/2} dW_6. \tag{18}$$

$$dx_7 = x_1/x_2 \cdot (x_6 - x_7)dt. \tag{19}$$

$$dx_8 = \{(q_2/x_2 - K_{81}) \cdot (x_7 - x_8) - K_{82} \cdot x_{11} + \delta_8\}dt. \tag{20}$$

$$dx_0 = (\alpha_9 + 1 - \beta_9 \cdot x_0)dt + (2\gamma_9 \cdot x_0)^{1/2} dW_9. \tag{21}$$

$$x_9 = 482.5/(32.6 + x_0). \tag{22}$$

$$dx_{10} = \{x_1/x_2 \cdot (x_9 - x_{10})\}dt. \tag{23}$$

$$dx_{11} = \{q_2/x_2 \cdot (x_{10} - x_{11}) + K_{82} \cdot (K_{11,1} - x_{11}) - K_{81} \cdot x_8 + K_{11,2} + \delta_{11}\}dt. \tag{24}$$

So far, a complete water pollution model for the non-tidal river system is obtained.

Combining the advantages of Harris' model, the three-steps model has the advantages of structure and information, such as applying the mass balance law, a general form of CSRT (continuously stirred reaction tank), taking the advantages of SDEs, fixing the three-step strategy with three SDEs representation, describing a separate but interacting relationship of material mixing processes and biochemical reaction processes.

For the benefit of comparing the input-output model (Harris' model), the equations of Harris' model are rewritten as follows:

$$dy_1 = (\alpha_1 + 1 - \beta_1 \cdot y_1)dt + (2\gamma_1 \cdot y_1)^{1/2} dW_1. \tag{25}$$

$$dy_2 = (y_1 - q_1)dt. \tag{26}$$

$$dy_3 = (\alpha_3 + 1 - \beta_3 \cdot y_3)/V dt + (2\gamma_3 \cdot y_3/y_1)^{1/2} dW_3. \tag{27}$$

$$dy_5 = \{q_2 \cdot (y_3 - y_5) - K_{51} \cdot y_2 \cdot [K_{b2} \cdot q_2 + K_{53} \cdot (y_3 - y_5)]\}/y_2 dt. \tag{28}$$

$$dy_6 = (\alpha_6 + 1 - \beta_6 \cdot y_6)/V dt + (2\gamma_6 \cdot y_6/y_2)^{1/2} dW_6. \tag{29}$$

$$dy_8 = \{(q_2/y_2 - K_{81}) \cdot (y_6 - y_8) - K_{82} \cdot y_{11} + \delta_8\}dt. \tag{30}$$

$$dy_0 = (\alpha_9 + 1 - \beta_9 \cdot y_0)dt + (2\gamma_9 \cdot y_0)^{1/2} dW_9. \tag{31}$$

$$y_9 = 482.5/(32.6 + y_0). \tag{32}$$

$$dy_{11} = \{q_2/y_2 \cdot (y_9 - y_{11}) + K_{82} \cdot (K_{11,1} - y_{11}) - K_{81} \cdot y_8 + K_{11,2} + \delta_{11}\}dt. \tag{33}$$

5 Estimated parameters and initial values

Some of these parameters were estimated by Young and Beck (1974) from observations taken over an extensive period on the River Cam, and some of these parameters refer to the Henry's law. The concentration of DO and BOD are followed by Harris' estimation.

The set of estimated parameters and initial values are assumed as following:
$\{x_0, \quad x_1, \quad x_2, \quad x_3, \quad x_4, \quad x_5, \quad x_6, \quad x_7, \quad x_8, \quad x_9, \quad x_{10}, \quad x_{11}\}$
$= \{20.0 centigrade; \quad 1.8 \cdot 10^5 liter/hour; \quad 5.5 \cdot 10^5 liter; \quad 29.996040 mg/l;$
$29.996040 mg/l; \quad 29.996040 mg/l; \quad 34.953805 mg/l; \quad 34.953805 mg/l;$
$34.953805 mg/l; \quad 9.187857 mg/l; \quad 3.195776 mg/l; \quad 3.195776 mg/l\};$

$\gamma_i = 0.35/1.5; \quad \beta_i = 0.35; \quad \alpha_i = x_{i,0} \cdot \beta_i - 1$
$K_{51} = 6.0 \cdot 10^{-3}/hour; \quad K_{b2} = 62.916849 mg/l; \quad K_{53} = 0.35;$
$K_{11,1} = 9.187857 mg/l; \quad K_{11,2} = 1.098548 mg/l/day;$
$\delta_8 = 0.0$ and $\delta_{11} = 0.0$,
$K_{82} = 0.3 \cdot 1.02^{T-20}/day$, by Henry's law,
$K_{81} = 0.17 \cdot 1.047^{T-20}/day$, by Henry's law.

6 Numerical models and its algorithms

6.1 Various methods of the numerical solution of SDEs

a) Monte Carlo method. It is somewhat inefficient because it does not use the special structure of these equations, specifically they are characterized by their drift and diffusion coefficients.

b) Finite state Markov Chains approach. It seems applicable only for low dimensional problems on bounded domains [10].

c) Kalman Filters. Estimation theory in the form of linear and extended Kalman filters have been used successfully in probabilistic water pollution models by several investigators including Beck and Young (1976), Bowles and Grenney (1978), and Bowles and Beck (1979). These models are useful for real time control and forecasting, but since they require contemporaneous measurements on the river system, they are not applicable for prediction of system performance under alternative control strategies for which it is not possible to obtain measurements.

d) Discrete time numerical methods [11].

An applicable approach to solving SDEs seems to be the simulation of sample paths of time discrete approximations. This is based on a finite discretization of the time interval [0,T] under consideration and generates approximate values of the sample paths, step by step, at the discretization times.

The simulated sample paths can then be analyzed by the usual statistical methods to determine how good the approximation is and in what sense it is close to the exact solution.

An advantage of considerable practical importance using this approach is that the computational costs such as time and memory required increase only polynomially with the dimension of the problem. Variance reduction methods allow a considerable decrease in the required sample size [11].

6.2 Some schemes of the discrete time numerical methods

a) The Euler scheme (see Maruyama (1955)).

b) The Runge-Kutta schemes.

c) The Stratonovich-Taylor expansions schemes.

d) The Ito-Taylor expansions schemes.

Early simulation studies and theoretical investigations by Clements and Anderson (1973), Wright (1974), Clark and Cameron (1980) and others showed that not all heuristic discrete time approximations of an SDE converge in a useful sense to the solution process as the maximum step size δ tends to zero.

In particular, it was found that one cannot simply use a deterministic numerical method for ordinary differential equations, such as a higher order Runge-Kutta method or Euler methods, which ignore the special properties of an SDE. (see appendix.)

In this paper, the Ito-Taylor expansions schemes would be applied.

As a sample of simulation of water pollution models, a 1.5 order Ito-Taylor expansion scheme is applied, but it does not go on to discuss which method is the best one, such as Stratonovich-Taylor scheme or Ito-Taylor-Runge-Kutta scheme.

The notation "1.5 order" is just followed by the linear notation, for the nonlinear case, the accurate convergence rate has not been proved.

In detail, refer to section 7.

7 The convergence of Numerical Solutions

In this section we consider numerical solutions for the following Stochastic Differential equation based on Ito-Taylor expansions:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \tag{34}$$

The equation is, here, interpreted as a stochastic integral equation

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)ds + \int_{t_0}^{t} b(s, X_s)dW_s \tag{35}$$

where the second integral is an Ito integral. The conditions for $a(t, X), b(t, X)$ will be stated later.

Let Y^δ be a general time discrete approximation with maximum step size δ. We say that Y^δ converges strongly to X at time T if

$$\lim_{\delta\to 0} E(|X_T - Y^\delta|) = 0 \tag{36}$$

We say Y^δ converges strongly with order γ at time T if there exists a positive constant C and δ_0 such that

$$\varepsilon(\delta) = E(|X_T - Y^\delta|) \leq C\delta^\gamma \tag{37}$$

for each δ with $0 < \delta < \delta_0$.

For a multi-dimensional Wiener process $W = (W^1, W^2, ...W^m)$, we define the multiple Wiener integrals by

$$I_{(j_1,j_2)} = \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^{\tau_{s_2}} dW_{s_1}^{j_1} dW_{s_2}^{j_2} \tag{38}$$

for $j_1, j_2 \in \{1, ..., m\}$ with $j_1 \neq j_2$, which cannot be expressed simply in terms of the increments $\Delta W_n^{j_1}$ and $\Delta W_n^{j_2}$ of the corresponding Wiener process.

In the general multi-dimensional case with $m = 1, 2, ...$ the k-th component of the order 1.5 strong Taylor scheme takes the form:

$$\begin{aligned}
Y_{n+1}^k &= Y_n^k + a\Delta + \frac{1}{2}L^0 a^k \Delta^2 && (39)\\
&+ \sum_{j=1}^{m}(b_{k,j}\Delta W^j + L^0 b^{k,j} I_{(0,j)+} + L^j a^k I_{(j,0)}) && (40)\\
&+ \sum_{j_1,j_2=1}^{m}(L^{j_1} b^{k,j_2} I_{(j_1,j_2)}) + \sum_{j_1,j_2,j_3=1}^{m}(L^{j_1} L^{j_2} b^{k,j_3} I_{(j_1,j_2,j_3)}). && (41)
\end{aligned}$$

where, the operators:

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d}(a^k \frac{\partial}{\partial y^k}) + \frac{1}{2}\sum_{k,l=1}^{d}\sum_{j=1}^{m} b^{k,j} b^{l,j} \frac{\partial^2}{\partial y^k \partial y^l}$$

$$L^j = \sum_{k=1}^{d}(b^{k,j}\frac{\partial}{\partial y^k})$$

Here ,we also have multiple Ito integrals with respect to different components of the Wiener process.

Let $\xi_j, \zeta_{j,1}, ..., \zeta_{j,p}, \eta_{j,1}, ..., \eta_{j,p}, \mu_{j,p}$ and $\phi_{j,p}$ be independent standard Gaussian random variables. Then for $j, j_1, j_2, j_3 = 1, 2, ..., m$ and $p = 1, 2, ...$ we have

$$I_{(j)} = \Delta W^j = \sqrt{\Delta}\xi_j,$$

$$I_{(j,0)} = \frac{1}{2}\Delta(\sqrt{\Delta}\xi_j + a_{j,0})$$

with $a_{j,0} = -\frac{\sqrt{2\Delta}}{\pi}\sum_{r=1}^{p}\frac{1}{r}\zeta_{j,r} - 2\sqrt{\Delta\rho_p}\mu_{j,p}$
where $\rho_p = \frac{1}{12} - \frac{1}{2\pi^2}\sum_{r=1}^{p}\frac{1}{r^2}$

$$I_{(0,j)} = \Delta W^j \Delta - I_{(j,0)},$$

$$I_{(j,j)} = \frac{1}{2}\{(\Delta W^j)^2 - \Delta\}$$

$$I^p_{(j_1,j_2)} = \frac{1}{2}\Delta\xi_{j_1}\xi_{j_2} - \frac{1}{2}\sqrt{\Delta}(\xi_{j_1}a_{j_2,0} - \xi_{j_2}a_{j_1,0} + A^p_{j_1,j_2}\Delta$$

for $j_1 \neq j_2$ where $A^p_{j_1,j_2} = \frac{1}{2\pi}\sum_{r=1}^{p}(\frac{1}{r}(\zeta_{j_1,r}\eta_{j_2,r} - \zeta_{j_2,r}\eta_{j_1,r}))$;
And we have $E(|I^p_\alpha - I_\alpha|^2) \leq C\frac{\Delta^2}{p}$ for the above approximations of the multiple Ito ingtegrals.

We need to choose p so that $p = p(\Delta) \geq \frac{K}{\Delta^2}$ for an appropriate positive constant K to ensure that the Ito-Taylor scheme (39) with these approximations of the multiple Ito integrals really does have order 1.5 of strong convergence for is a autonomous equations.

The algorithm of simulation of our water pollution model does not present a $I_{(j_1,j_2,j_3)}$ term because its coefficients equal zero, so we skip it here.
(see [11] 350-357)

We apply the above numerical methods to our water polution models and prove the strong convergence for them:

Theorem 7.1 *Suppose that the sludge flow q_0 and C_{BOD0} is bouded below,*

$$q_0 \geq Q, \quad C_{BOD0} \geq C \tag{42}$$

for some positive constant Q and C. Then our scheme of numerical methods converges strongly to the solution of SDE with order 1.5.

proof
The crucial point is to estimate the *jump* of random number. In our case, the critical terms are

$$(2\gamma_6 \cdot c_{BOD0}/V)^{1/2}dW_6, \quad (2\gamma_1 \cdot q_0)^{1/2}dW_1$$

From the assumptions in the theorem, we note that the absolute values of the differentials of the functions $(2\gamma_6 \cdot c_{BOD0}/V)^{1/2}, (2\gamma_1 \cdot q_0)^{1/2}$ are bounded.

Therefore all the coefficients in our SDEs satisfy both the Lipshitz condition and the linear growth bound:

$$|a(t,x) - a(t,y)| \leq K|x-y|$$
$$|b(t,x) - b(t,y)| \leq K|x-y|$$
$$|a(t,x)^2| \leq K^2(1+|x|^2)$$
$$|b(t,x)^2| \leq K^2(1+|x|^2)$$

Under these conditions we can conclude that our numerical solution converges strongly to the solution of SDEs. For the order 1.5, see the section 10.4 in [11].

8 Simulation results

By 1.5 order Ito-Taylor expansions schemes, we summarize our simulation results as follow:

1) The theory of SDEs is a beneficial tool for studying diffusion process and system errors accounting. It can describe the physical relationships of different factors (e.g. BOD gain, in figure 5), and can predict phenomenon on the environment by simulation and statistical analysis (e.g. figure 6,7,8,).

2) By adding a mixing process to the input/output model, the three-steps model divides the system into mixing process (stirring) and biochemical process, exploits the information of pre-biochemical reaction on the system, so it increases the accuracy of model representing the real physical relationships (It is based on the criterion of models: the quantity of fluid mixing. M.B.Beck).

3). Figure 3 shows that the sample paths of the output pollutant concentrations between the two models are almost identical, but figure 4 shows the sample paths of the output BOD concentrations and the output DO concentrations in each model are different. This means there are some potential different advantages for each model.

4) Figure 11, 12 and 13 show that the output DO concentration of the three-steps model seems to be determined by the input DO concentration, namely it relates closely to the temperature here (because the input DO concentration is assumed to be determined by temperature only), but Harris' model shows that it is related to the input BOD concentration.

9 Discussion

9.1 The relationship between the three-steps model and SDEs

The primary significances of the three-steps model contrasting SDEs and diffusion processes are:

1) It suits the spirit of the Ito process which represents general diffusion processes as transformations of a Brownian motion (Michael Turelli. (1976)), either in discrete time cases or continuous time cases;

2) Roughly speaking, it suits the idea that a sample of a stochastic process is continuous/deterministic in a sufficient small time interval dt, so a stable state need to be assumed as the basic starting point for modeling in the time interval dt;

3) It makes clear what a physical diffusion process means, since the diffusion process it represents depends on which interpretation is employed.

4) It can single out the different diffusion processes by different Wiener processes.

On the other hand, a Wiener process in SDEs can also be used to account for the system errors only, that means we can lump the different diffusion processes/noises all together as one noise term, rather than distinguish the more useful information concerning the effects of environmental stochasticity.

9.2 Limits of the three-steps model

Determining the degree and structure of the interaction between variables and model components is difficult to quantify (see section 1, or (R.H.McCuen)). A comparison of sensitivity estimates may provide some indications of the degree of interaction, but determining the quality of a model must often be based on many other theoretical considerations.

a) Though we show some advantages of the three-steps model, it must be verified by the observed data. It is an important step but is lacking at present in this paper.

b) It looks like the three-steps model is logically completed, but the equation of describing mixing process and the equation of describing reaction process are related to each other. From a mathematical sense, they are not independent. Therefore theoretically speaking, at this time, some of the mathematical principles would be in conflict with the new model in reality. On the other hand, it can be performed by digital computer method. (e.g. the value of $c_{DO1}(= c_{DO2})$ and c_{DO3} are exchanged, step by step, just like a cycle during programming), this due to the advantage of digital techniques.

c) Because the equations are non-linear SDEs, usually, we cannot obtain the explicit solutions of SDEs, but we can get a numerical solution. In a non-linear system, how to determine the confidence interval theoretically becomes a problem.

9.3 On improving and calibration of the three-steps model

There are some techniques for improving the three-steps model and calibrating the three steps model.

1) If the noise components are state-correlated or time-series-correlated, we can deal with them by using another mathematic technique.

a) In general, the different factors (e.g. the influent feed-rate and substrate concentration) are related. They are reflected in the model (e.g. equation (1) and (3)) by the correlated Wiener processes (e.g. W_1 and W_3). We can deal with them by mathematical techniques (see Di Toro and Van Straten (1979), or Harris (1975)).

b) If the noises are time-series-correlated, (e.g. $W_1(t_i)$ and $W_3(t_j)$ are not iid. (independent, identically distributed)), then the solutions would be very complex, such as Ito's solution, Stratonovich's solution, or neither of them. (see T.E.Unny (1984) or Michael Turelli (1976)).

2) A more general model can be presented by adding some noise terms to the reacting equations respectively, e.g. x_5 or x_8 or x_{11}, to account for poor modeling as well as additive disturbances (e.g. effects of sunlight on photosynthesis of algae or wind gusts to account for the turbulent reaction process). It is possible to improve the model so that it can become a more reasonable mathematical approximation of the physical biochemical exchanges occurring in the river system.

3) This input/mixing/output structure of the model can be applied to other fields of study, such as a polluted lake, or sea, etc. It is a form of the continuously stirred tank reactors (CSTR).

4) If we relate the ideas of the Ito integral method and Stratonovich integral method with Harris' model and the three-steps model, we find that Harris' model and the three-steps model correspond to the ideas of the Ito integral and the Stratonovich integral respectively, because in Harris' model, the value of reaction concentrations are evaluated at instant time t_n in $[t_n, t_{n+1}]$, but in the three-steps model, they are accounted for by the average concentrations on time interval $[t_n, t_{n+1}]$ and space volume V. In this view, we cannot say which model is better.

10 Conclusion

In this study, we have investigated two water pollution models based on SDEs, (the input/output model (Harris' model) and the input/mixing/output model (the three-steps model)).

The results obtained in this study are summarized as following:

1) The theory of SDE is a beneficial tool for studying water pollution. Based on the analysis of computer simulation of SDEs, we can obtain predictions on phenomenon, and the relationship between different factors considered. From the predictions and the relationships, we can decide how to control and plan political strategies for self-purification

river systems.

2) Comparing the ODEs (ordinary differential equations) dynamic model, an additional term of the Wiener process in the SDEs model can be used to account for the system errors only, or to represent the diffusion process.

3) The three-steps model is improved from the input/output model by adding a mixing (stirring) process. It exploits the information of pre-biochemical reaction on the system. From the model aspect, it represents the system more efficiently than the input/output model, but it is more difficult to apply. Respectively, they have some potential different advantages .

4) If the term of prediction is short, e.g. 5 days, both of the models are proper to predict the phenomena; if the term of prediction is a little longer, e.g. 6-10 days, the three-steps model has an advantage over the input/output model; but if the term of prediction is long, both of the models are somewhat improper to predict the future, at the time, the three-steps model loses the advantage of predicting BOD.

5) In the three-steps model, output DO concentration seems to be determined by the input DO concentration, but in the input/output model, it seems to be determined by the input BOD concentration.

Further work: The determination of the quality of a model must be based on many other theoretical considerations, e.g. it must be verified by the observed data; in a non-linear system; the methods of how to determine a confidence interval theoretically must be developed; the ideas of how to evaluate the concentrations of material reaction in the input/output model and the three-steps model, are similar to the Ito integral method and Stratonovich integral method respectively. What are the relationships between them?

11 References

1. C.J.Harris. (1975), 'simulation of non-linear stochastic equations with applications in modeling water pollution', Mathematical models for environmental problems, 269-281.
2. R.H.McCuen. (1975),'The anatomy of the modeling process', Mathematical models for environmental problems, 401-412.
3. C.J.Harris. (1979),'River quality models near effluent outfall', Mathematical modelling of turbulent diffusion in the environment, 459-496.
4. C.J.S.Petrie. (1978),'The modeling of engineering systems-mathematical and computational techniques', Mathematical models in water pollution control, 137-163.
5. M.B.Beck. (1978),'Modelling of dissolved oxygen in a non-tidal stream', mathematical models in water pollution control, 137-163.
6. T.E.Unny. (1984), 'Numerical Integration of stochastic differential equations in catchments modeling', Water resources research, Vol.20,NO.3, 369-369.
7. Michael Turelli. (1977), 'Random environments and stochastic calculus', Theory pop. biology, Vol.12. 140-178.
8. Peter Young and Bruce Beck. (1974), 'The modeling and control of water quality in river system ', Automatic Vol.10, 455-468.
9. D.Barnes et. al. (1981) ,'water and wastewater engineering system', Pitman.
10. Brad A.Finney et. al. (1982), Random differential equations in river water quality modeling, Water resources research, Vol.18, NO.1, 122-134.
11. Peter E.Kloeden and Echhard Platen, (1992), 'Stochastic numerical methods', Springer-Verlag.
12. Willian H.Press et al. 1988, 'numerical Recipes in C', http://www.nr.com/

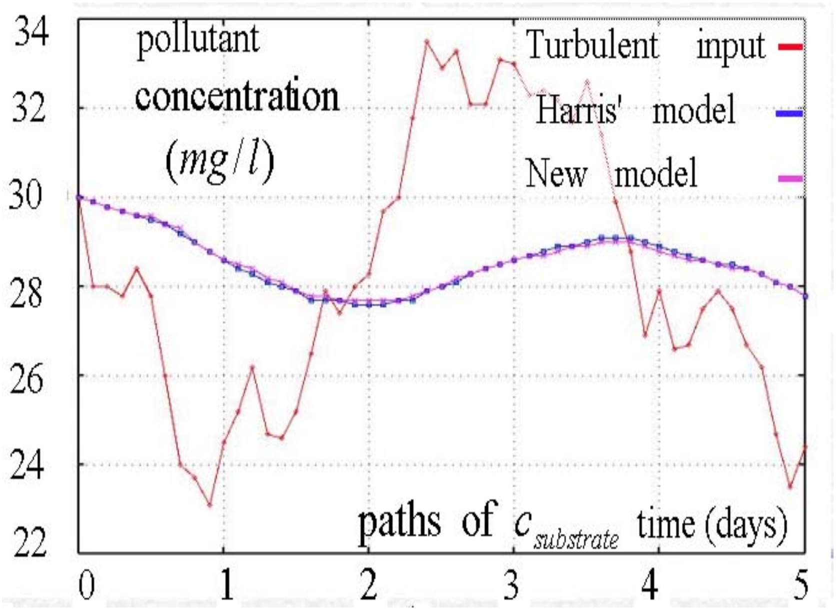

Figure 3: paths of substrate concentrations

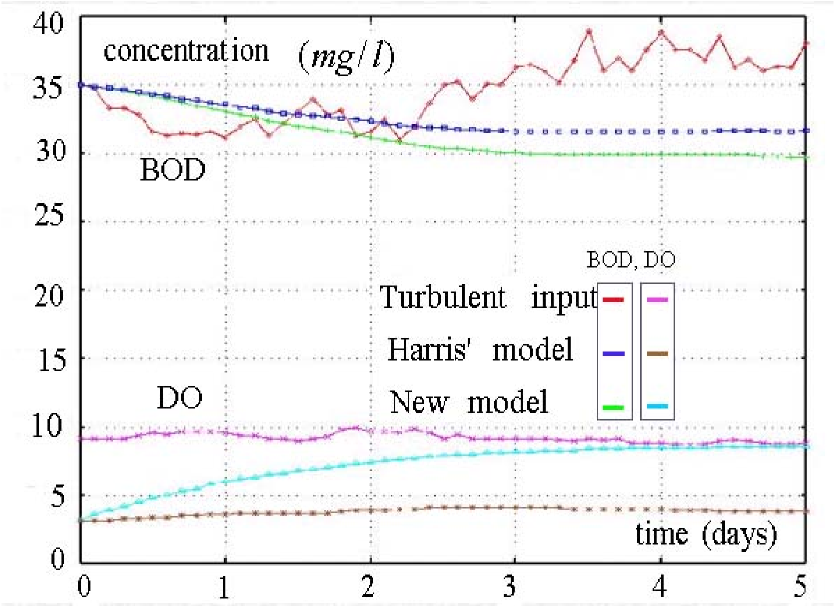

Figure 4: paths of BOD and DO concentrations

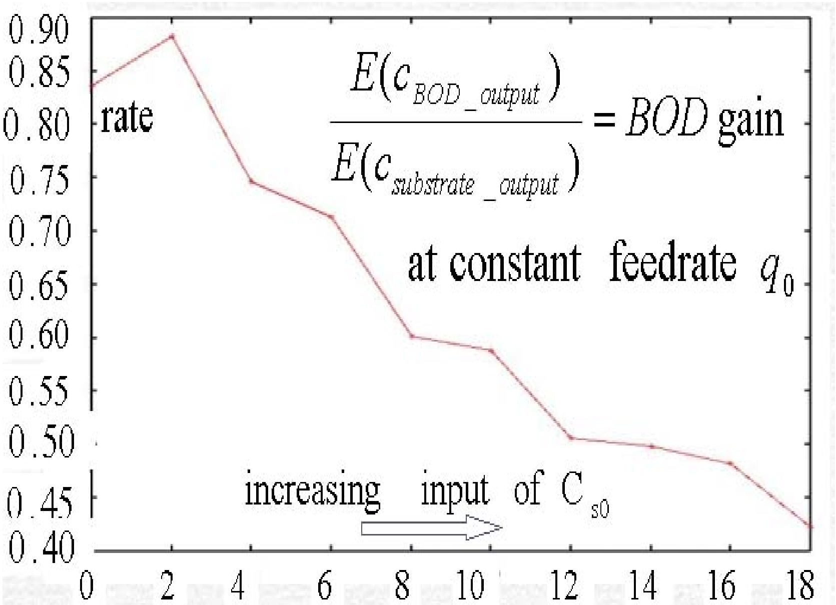

Figure 5: BOD gain; $unit = E(c_{s0}^{(2)})/(\beta_3 E(c_{s0}))$

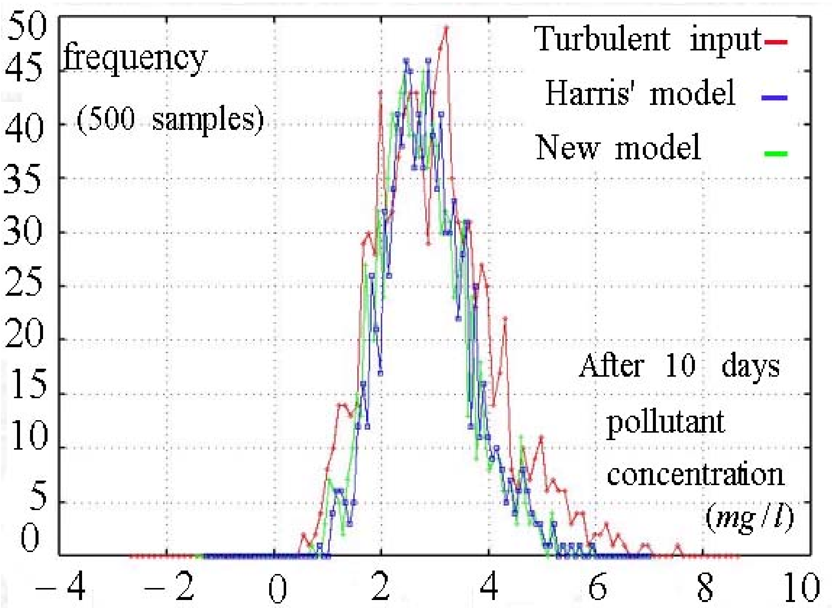

Figure 6: Distributions of substrate concentrations

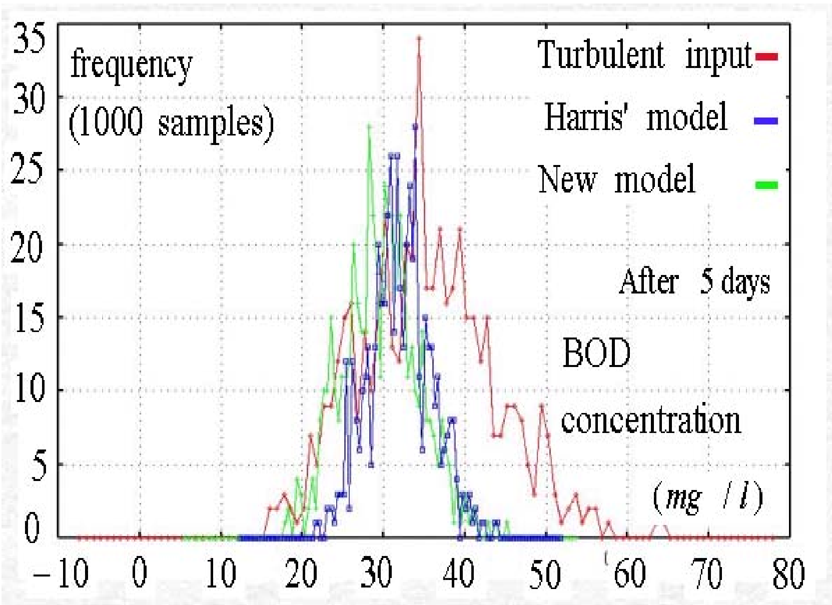

Figure 7: Distributions of BOD concentrations

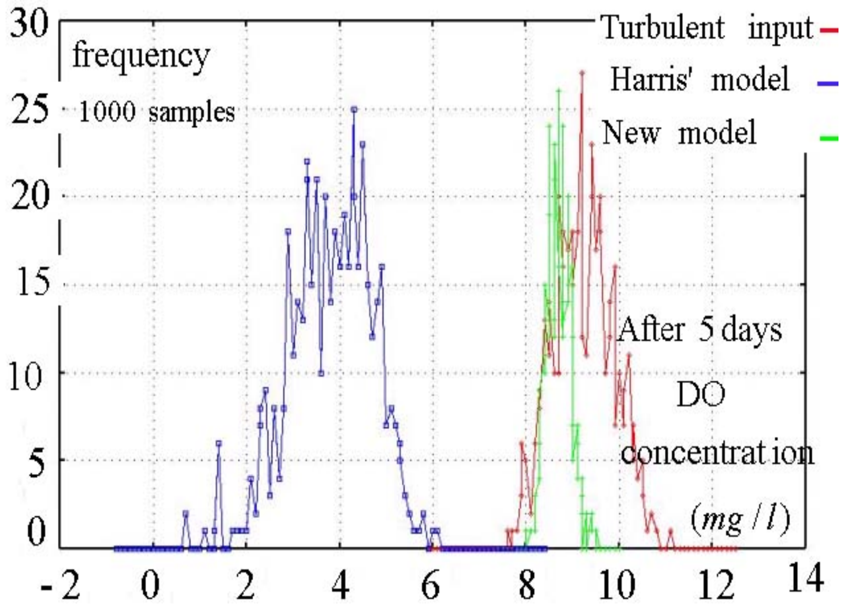

Figure 8: Distributions of DO concentrations

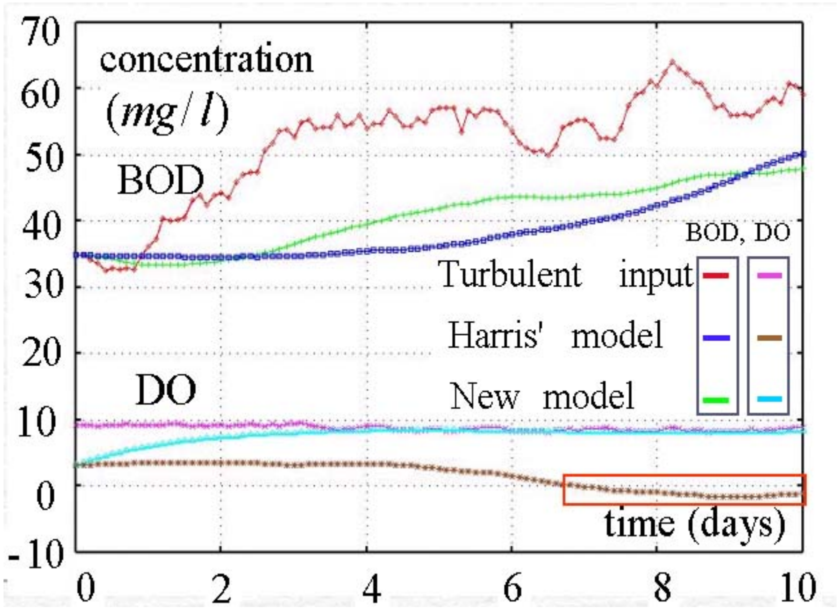

Figure 9: Paths of DO concentrations go down to zero

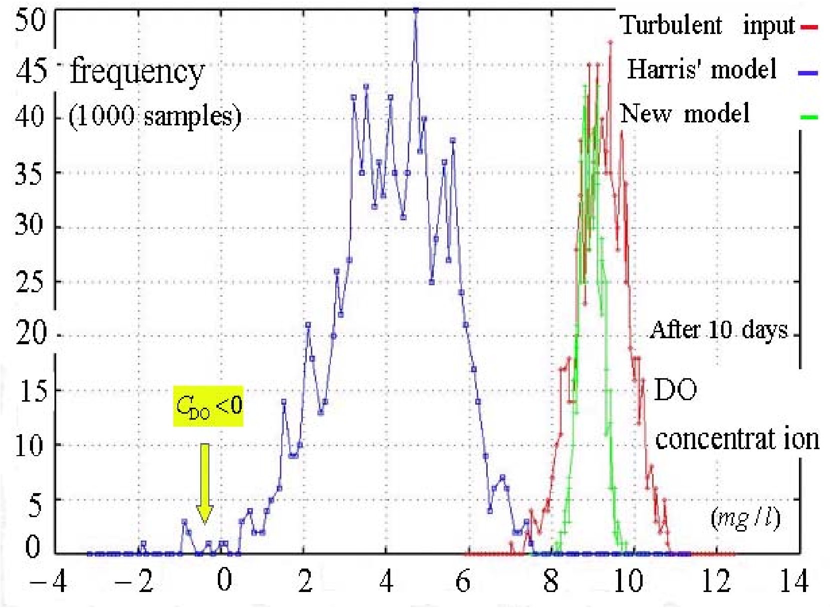

Figure 10: Distribution of DO concentrations including negative samples

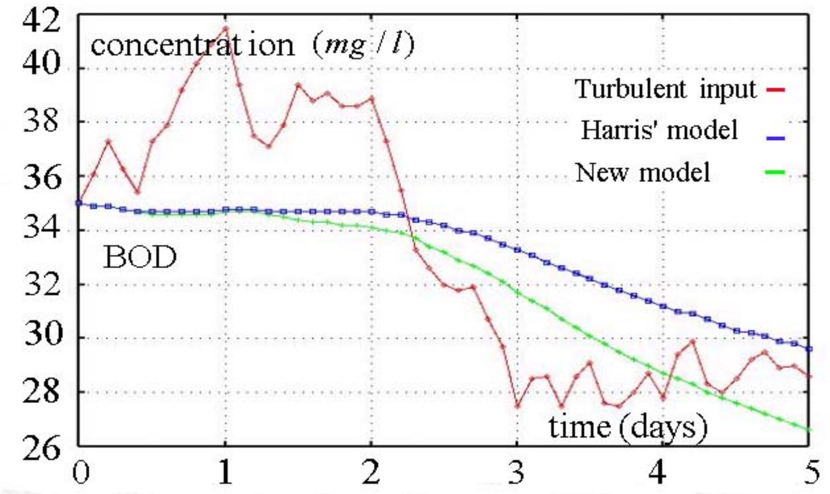

Figure 11: paths of BOD concentrations

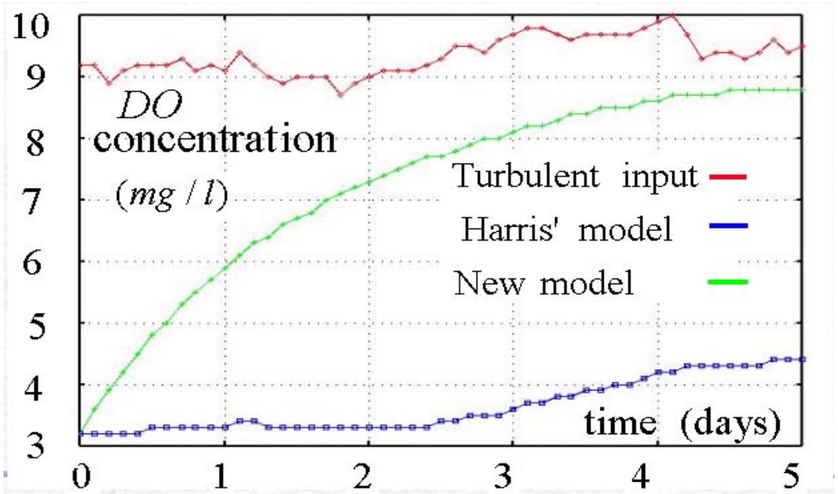

Figure 12: paths of DO concentrations

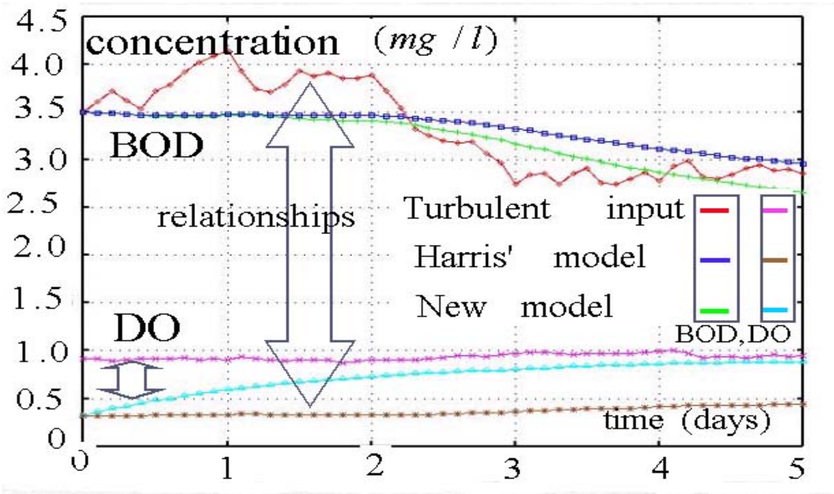

Figure 13: paths of BOD and DO concentrations