East-West J. of Mathematics: Vol. 23, No 2 (2022) pp. 138-154
https://doi.org/10.36853/ewjm0391

FORWARD PROBLEM FOR POLLUTION CONTROL BASED ON STREETER-PHELPS EQUATION

Tran Thi Nang, Nguyen Dinh Huy and **Bui Ta Long**

Department of Mathematics, Faculty of Fundamental Sciences
Van Lang University, Ho Chi Minh City, Vietnam
e-mail: nang.tt@vlu.edu.vn

†*Department of Applied Mathematics, Faculty of Appied Sciences,*
Ho Chi Minh City University of Technology,
National Univ. of Ho Chi Minh City, Ho Chi Minh City, Vietnam
e-mail: dinhhuy56@hcmut.edu.vn

‡*Laboratory for Environmental Modelling,*
Faculty of Environment and Natural Resources
Ho Chi Minh City University of Technology
National Univ. of Ho Chi Minh City, Ho Chi Minh City, Vietnam
e-mail: longbt62@hcmut.edu.vn

Abstract

The relationship between the discharger and the water quality of the receptor streams was described by a mathematical equation, for the first time introduced in the early 20th century, known as the Streeter – Phelps equation. There have been many follow-up studies to develop this work; however, most of these stopped in the case of steady-state flows and discharges. The expansion through the unsteady case is unnoticed. In this study, the Streeter-Phelps equation is considered in its most general form, accounting for the temporal variation of both the flow and the discharger over time. This study inherits the previous studies but considers for the case of an unstead discharger, especially when the pollutant concentration in the wastewater exceeds the allowable standard. The mathematical model Streeter – Phelps is applied in the case of insteady of the waste discharger. The results show that the error with

Key words: Crank – Nicolson scheme, discharger - receptor, MIKE, simulation error, simulation experiment, water quality modelling.
2010 AMS Classification:

other products like MIKE is in the range of less than 10%. This result allows the application of the Streeter - Phelps model to the pollution control problem.

1 Introduction

The year 1925 marked the birth of the first scientific work in the field of water quality modelling, in which Streeter - Phelps (SP) successfully built a mathematical relationship showing the relationship between the discharger and the quality of river water receiving the waste (the receptor) [1]. Subsequent developments were presented in a series of scientific publications in the following three directions: developing a method to calculate the reaeration coefficient due to different natural flows [2] - [4]; developing mathematical models that pay attention to the impact of hydrodynamic and hydrological factors on pollutant transport [5]; and developing computer programs to calculate the influence of waste dischargers on stream water quality [6] - [10].

Within the framework of the mentioned studies, the "forward" problem was posed, and some analytical solutions were presented. Based on discharge characteristics, the concentration of discharging pollutants is determined [11] - [14]. In the case of unsteady flow, the methods used in [1], [12] – [13] did not work. Therefore, it is necessary to develop another alternative approach, specifically to find the solution of the forward problem based on the numerical solution of the SP equation (hereafter called the Streeter-Phelps model, denoted by SPM) [15].

Due to the current difficulties in continuous monitoring technology for unsteady model, the actual measured dataset was replaced by the MIKE running dataset - one of the reputable products exported. The [15] - [17] noted that the real time series dataset in practice is not easy to find, and thus, the real measured dataset was replaced by the modelling results using the MIKE tool. Note that the same idea in the field of engineering is expressed in [16].

Extending the results of the SPM studies for the steady to the unsteady case is not only theoretically meaningful, but will help solve the water pollution control problem. In the event of an emergency, some wastesources may discharge exceed the required capacity leading to pollution accident situations. The incident damage needs to be quantified that leading to using unsteady SPM. Since then, the given work has the aim of building and verifying the numerical solution scheme and its application for a case study. The results are not only limited to numerical solutions but also help to calculate the relationship "waste source - receptor" served effectively the problem of environmental protection and sustainable development.

2 Problem description

Water quality modelling of rivers and streams goes back to 1925 and the Streeter – Phelps equation for the dissolved oxygen concentration in the Ohio River [1]. There have been many developments since that time that have made it possible to model other settings in greater detail and with more realistic processes [18]. Modelling the aquatic chemistry of BOD, TSS, $N - NH_4^+$ and dissolved oxygen in natural waters requires mathematical formulations for mass transport in a stream or a river and reaction kinetics.
For the case of time – variations in BOD in a river of estuary with constant freshwater flowrate, cross – sectional area, and dispersion coefficient, the formula

$$DO = DO_{saturated} - D(x) \tag{1}$$

will determine the dissolved oxygen concentration at position x, that is, determine the extent of the influence on the river due to the impact of the waste discharger. The DOsaturated parameter is calculated based on the table of DO values at integer temperatures and is interpolated at other temperature values [18].
It is a parabolic, partial differential mass balance equation. The general form of the equation is given as

$$\frac{\partial L}{\partial t} = -\frac{1}{A}\frac{\partial (QL)}{\partial x} + \frac{1}{A}\frac{\partial}{\partial x}\left(EA\frac{\partial L}{\partial x}\right) - k_1 L \tag{2}$$

in which the cross-sectional area, flow rate, and dispersion coefficient can vary in time and space. The DO deficit equation is developed in a similar manner:

$$\frac{\partial D}{\partial t} = -\frac{1}{A}\frac{\partial (QD)}{\partial x} + \frac{1}{A}\frac{\partial}{\partial x}\left(EA\frac{\partial D}{\partial x}\right) + k_2 L - k_1 D \tag{3}$$

In the case that E, Q, and A in a certain period of time do not change, then the form (2) and (3) can be written as

$$\begin{cases} \frac{\partial L}{\partial t} + u\frac{\partial L}{\partial x} = E\frac{\partial^2 L}{\partial x^2} - k_1 L \\ \frac{\partial D}{\partial t} + u\frac{\partial D}{\partial x} = E\frac{\partial^2 D}{\partial x^2} + k_1 L - k_2 D \end{cases} \tag{4}$$

To be able to model the change in the time-varying BOD concentration and the DO deficit in space and time, it is necessary to set the initial and boundary conditions:

$$\begin{cases} L(t,x)|_{x=0} = L_0(t);\, L(t,x)|_{t=0} = L_1(x) \\ D(t,x)|_{x=0} = D_0(t);\, D(t,x)|_{t=0} = D_1(x) \\ x \in [0;L]; t \in [0;T] \end{cases} \tag{5}$$

where $u(m/s)$ is the stream velocity, and it is a piecewise variable function; $L(x,t)$ (mg/l) is the amount of oxidizable organic material as oxygen equivalents (abbreviated as BOD); $D(x,t)(mg/l)$ is the DO deficit identified from (1); $x(m)$ is the distance along the reach moving downstream; $k_1(day^{-1})$ is the decomposition rate in the stream; $k_2(day^{-1})$ is the surface reaeration rate; and $E(m^2/s)$ is the dispersion coefficient. Equations (3) and (4) show the relationship between the "Discharger – Receptor", in which the "dirty" level of the waste discharger is expressed through parameter $L_0(t)$ = total BOD quantity from the waste discharger, which was placed at point x_0; $L_1(x)$ = BOD background concentration; and $D(x,t)$ = dissolved oxygen deficiency (mg/l) at position $x(m)$ from the source represents the quality of the river water affected by the waste discharger. $D_0(t) = DO$ deficit identified at cross Section x_0, and $D_1(x) = DO$ deficit background concentration on $[0;L]$ [15].
To find the solution of the forward problem based on the numerical solution of the SP equation.

3 Methods and data

3.1 Crank–Nicolson-type schemes

Divide the stream segment [0;L] into N equal reaches, the distance between the two adjacent locations is denoted by Δx. Divide the time interval [0;T] into M equal segments, each of which is Δt. In plane Oxy, there is a rectangular difference grid D:
$D = (x_i, t_j); x_i = i\Delta x, t_j = j\Delta t; i = \overline{1;N}; j = \overline{1;M}$. Set $\sigma_{i,n} = u_{i,n}\frac{\Delta t}{\Delta x}$, $\delta_i = E_i\frac{\Delta t}{(\Delta x)^2}$
At the node at position i=1:

$$\frac{L_1^{n+1} - L_1^n}{\Delta t} + u_{1,n}\frac{L_2^n - L_0^n}{2\Delta x} = E_1\left(\frac{L_2^n - 2L_1^n + L_0^n}{\Delta x^2}\right)$$
$$-k_1\left[\theta_2 L_2^n + (1+\theta_2)L_0^n\right]$$

$$\Rightarrow L_1^{n+1} = \left[\frac{\sigma_{1,n}}{2} + \delta_1 - k_1\Delta(1-\theta_2)\right]L_0^n + (1-2\delta_1)L_1^n$$
$$-\left[\frac{\sigma_{1,n}}{2} - \delta_1 + k_1\Delta\theta_2\right]L_2^n$$

At the i-th position node with $2 \le i \le N-1$:

$$\frac{L_i^{n+1}-L_i^n}{\Delta t}+u_{i,n}\left\{\theta_1\frac{L_{i+1}^{n+1}-L_{i-1}^{n+1}}{2\Delta x}+(1+\theta_1)\frac{L_{i+1}^n-L_{i-1}^n}{2\Delta x}\right\}$$
$$=E_i\left\{\theta_1\frac{L_{i+1}^{n+1}-2L_i^{n+1}+L_{i-1}^{n+1}}{\Delta x^2}+(1+\theta_1)\frac{L_{i+1}^n-2L_i^n+L_{i-1}^n}{\Delta x^2}\right\}$$
$$-k_1\left[\theta_1\left\{\theta_2L_{i+1}^{n+1}+(1-\theta_2)L_{i-1}^{n+1}\right\}+(1-\theta_1)\left\{\theta_2L_{i+1}^n+(1-\theta_2)L_{i-1}^n\right\}\right]$$

$$\Rightarrow\left[\theta_1\left(-\frac{1}{2}\sigma_{i,n}-\delta_i\right)+\theta_1k_1\Delta t(1-\theta_2)\right]L_{i-1}^{n+1}+(1+2\theta_1\delta_i)L_i^{n+1}$$
$$+\theta_1\left(\frac{1}{2}\sigma_{i,n}+\theta_2k_1\Delta t-\delta_i\right)L_{i+1}^{n+1}$$
$$=\left[(1-\theta_1)\frac{\sigma_{i,n}}{2}+(1-\theta_1)\delta_i-(1-\theta_1)k_1(1-\theta_2)\Delta t\right]L_{i-1}^n+\left[1-2(1-\theta_1)\delta_i\right]L_i^n$$
$$\left[-(1-\theta_1)\frac{\sigma_{i,n}}{2}+(1-\theta_1)\delta_i-(1-\theta_1)k_1\theta_2\Delta t\right]L_{i+1}^n$$

At the node at position $i=N$:

$$\frac{L_N^{n+1}-L_N^n}{\Delta t}+u_{N,n}\left[\theta_1\frac{L_N^{n+1}-L_{N-1}^{n+1}}{\Delta x}+(1-\theta_1)\frac{L_N^n-L_{N-1}^n}{\Delta x}\right]$$
$$=E_N\left[\theta_1\left(\frac{L_N^{n+1}-2L_{N-1}^{n+1}+L_{N-2}^{n+1}}{\Delta x^2}\right)+(1-\theta_1)\left(\frac{L_N^n-2L_{N-1}^n+L_{N-2}^n}{\Delta x^2}\right)\right]$$
$$-k_1\left(\theta_1L_N^{n+1}+(1-\theta_1)L_N^n\right)$$

$$\Rightarrow-\theta_1\delta_NL_{N-2}^{n+1}+(2\theta_1\delta_N-\theta_1\sigma_{N,n})L_{N-1}^{n+1}+(1+\theta_1\sigma_{N,n}-\theta_1\delta_N+k_1\theta_1\Delta t)L_N^{n+1}$$
$$=(1-\theta_1)\delta_NL_{N-2}^n+\left[(1-\theta_1)\sigma_{N,n}-2(1-\theta_1)\delta_N\right]L_{N-1}^n$$
$$+\left[1-(1-\theta_1)\sigma_{N,n}+(1-\theta_1)\delta_N-k_1(1-\theta_1)\Delta t\right]L_N^n$$

The system of N equations is set up with the unknowns being the values of $L_1^{n+1},L_2^{n+1},\ldots,L_{N-1}^{n+1},L_N^{n+1}$ at the grid nodes $(i;j)$ with $j=n+1$, calculated based on the value of L at the position nodes at the previous time step.
The matrix form of the system of equations is $AL\times L^{n+1}=SL^n$, with the coefficient matrix AL being:

$$AL = \begin{bmatrix} AL_{P_1} & -AL_{E_1} & 0 & \dots & \dots & \dots & \dots \\ -AL_{W_2} & AL_{P_2} & -AL_{E_2} & 0 & \dots & \dots & \dots \\ 0 & -AL_{W_3} & AL_{P_3} & -AL_{E_3} & 0 & \dots & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \dots & \dots & \dots & 0 & -AL_{W_{N-1}} & AL_{P_{N-1}} & -AL_{E_{N-1}} \\ \dots & \dots & \dots & 0 & -AL_{WW_N} & -AL_{W_N} & AL_{P_N} \end{bmatrix}$$

With

$$\begin{aligned} &AL_{P_1} = 1, \\ &AL_{E_1} = 0 \\ &AL_{W_i} = \theta_1\left(\frac{1}{2}\sigma_{i,n} + \delta_i\right) - \theta_1 k_1 \Delta t(1-\theta_2), \\ &AL_{P_i} = 1 + 2\theta_1\delta_i, \\ &AL_{E_i} = -\theta_1\left(\frac{1}{2}\sigma_{i,n} + \theta_2 k_1\Delta t - \delta_i\right), 2 \le i \le N-1 \\ &AL_{WW_N} = \theta_1\delta_N, \\ &AL_{W_N} = -2\theta_1\delta_N + \theta_1\sigma_{N,n}, \\ &AL_{P_N} = 1 + \theta_1\sigma_{N,n} - \theta_1\delta_N + k_1\theta_1\Delta t \end{aligned}$$

and the freedom matrix is $SL^n = \begin{bmatrix} SL_1^n & SL_2^n & \dots & SL_{N-1}^n & SL_N^n \end{bmatrix}$.
In which

$$\begin{aligned} SL_1^n = &\left[\frac{\sigma_{1,n}}{2} + \delta_1 - k_1\Delta t(1-\theta_2)\right] L_0^n + (1-2\delta_1)L_1^n \\ &- \left[\frac{\sigma_{1,n}}{2} - \delta_1 + k_1\Delta t\theta_2\right] L_2^n \end{aligned}$$

$$\begin{gathered} SL_i^n = \left[(1-\theta_1)(\frac{\sigma_{i,n}}{2} + \delta_i) + (1-\theta_1)k_1(1-\theta_2)\Delta t\right] L_{i-1}^n \\ + \left[1 - 2(1-\theta_1)\delta_i\right] L_i^n + \left[-(1-\theta_1)\frac{\sigma_{i,n}}{2} + (1-\theta_1)\delta_i - (1-\theta_1)k_1\theta_2\Delta t\right] L_{i+1}^n, \\ 2 \le i \le N-1 \end{gathered}$$

$$\begin{aligned} SL_N^n = &(1-\theta_1)\delta_N L_{N-2}^n + \left[(1-\theta_1)\sigma_{N,n} - 2(1-\theta_1)\delta_N\right] L_{N-1}^n \\ &+ \left[1 - (1-\theta_1)(\sigma_{N,n} - \delta_N + k_1\Delta t)\right] L_N^n \end{aligned}$$

By solving this system of equations, we get the value of L at the (n+1)th time step. The values are substituted in the second equation in (4) to find D.
By applying the same difference formulas as when solving the first equation and

substituting the values of L at the (n+1)th time step, we obtain a system of N equations with the matrix form $AD \times D^{n+1} = SD^n$. The coefficient matrix AD is

$$AD = \begin{bmatrix} AD_{P_1} & -AD_{E_1} & 0 & \dots & \dots & \dots & \dots \\ -AD_{W_2} & AD_{P_2} & -AD_{E_2} & 0 & \dots & \dots & \dots \\ 0 & -AD_{W_3} & AD_{P_3} & -AD_{E_3} & 0 & \dots & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \dots & \dots & \dots & 0 & -AD_{W_{N-1}} & AD_{P_{N-1}} & -AD_{E_{N-1}} \\ \dots & \dots & \dots & 0 & -AD_{WW_N} & -AD_{W_N} & AD_{P_N} \end{bmatrix}$$

With

$$\begin{aligned} AD_{P_1} &= 1, \\ AD_{E_1} &= 0 \\ AD_{W_i} &= \theta_1 \left(\frac{1}{2}\sigma_{i,n} + \delta_i \right) - \theta_1 k_2 \Delta t (1-\theta_2), \\ AD_{P_i} &= 1 + 2\theta_1 \delta_i, \\ AD_{E_i} &= -\theta_1 \left(\frac{1}{2}\sigma_{i,n} + \theta_2 k_2 \Delta t - \delta_i \right), 2 \leq i \leq N-1 \\ AD_{WW_N} &= \theta_1 \delta_N, \\ AL_{W_N} &= -2\theta_1 \delta_N + \theta_1 \sigma_{N,n}, \\ AD_{P_N} &= 1 + \theta_1 \sigma_{N,n} - \theta_1 \delta_N + k_2 \theta_1 \Delta t \end{aligned}$$

and the freedom matrix is $SD^n = \begin{bmatrix} SD_1^n & SD_2^n & \dots & SD_{N-1}^n & SD_N^n \end{bmatrix}$ with

$$\begin{aligned} SD_1^n = &\left[\frac{\sigma_{1,n}}{2} + \delta_1 - k_2 \Delta t (1-\theta_2) \right] D_0^n + (1 - 2\delta_1) D_1^n \\ &- \left[\frac{\sigma_{1,n}}{2} - \delta_1 + k_2 \Delta t \theta_2 \right] D_2^n + k_1 \Delta t \left[\theta_2 L_2^n + (1-\theta_2) L_0^n \right] \end{aligned}$$

$$\begin{gathered} SD_i^n = \left[(1-\theta_1)(\frac{\sigma_{i,n}}{2} + \delta_i) + (1-\theta_1) k_2 (1-\theta_2) \Delta t \right] D_{i-1}^n + \left[1 - 2(1-\theta_1)\delta_i \right] D_i^n \\ + \left[(1-\theta_1)(\delta_i - \frac{\sigma_{i,n}}{2}) - (1-\theta_1) k_2 \theta_2 \Delta t \right] D_{i+1}^n \\ + k_1 \Delta t \left[\theta_1 \left\{ \theta_2 L_{i+1}^{n+1} + (1-\theta_2) L_{i-1}^{n+1} \right\} + (1-\theta_1) \left\{ \theta_2 L_{i+1}^n + (1-\theta_2) L_{i-1}^n \right\} \right], \\ 2 \leq i \leq N-1 \end{gathered}$$

$$\begin{aligned} SD_N^n = &(1-\theta_1)\delta_N D_{N-2}^n + \left[(1-\theta_1)\sigma_{N,n} - 2(1-\theta_1)\delta_N \right] D_{N-1}^n \\ &+ \left[1 - (1-\theta_1)(\sigma_{N,n} - \delta_N + k_2 \Delta t) \right] D_N^n \\ &+ k_1 \Delta t \left(\theta_1 L_N^{n+1} + (1-\theta_1) L_N^n \right) \end{aligned}$$

3.2 Dataset

3.2.1 Create a dataset for MIKE

Topographic data are extracted from general data across the basin, and the selected river segment is Tan Dinh stream, Binh Duong Province, Vietnam. The river segment has a length of L=1414 m, divided into 14 equally spaced reaches with cross sections numbered ID0 – ID14. The cross section and elevation are shown in Figure 1, down. To run the HD module, it is necessary to determine the discharge at ID0 and the water level at ID14. The discharge boundary data at the cross section are extracted from the NAM run results in the form of a time series with a time step of 1 h, the period from Jan 01, 2017 to Dec 31, 2017 [19] (Figure 2). The water level boundary is extracted from the Tide Prediction tool – a tool in MIKE 21. The extracted times coincide with the water level boundary times described above at cross section ID14 (Figure 3).

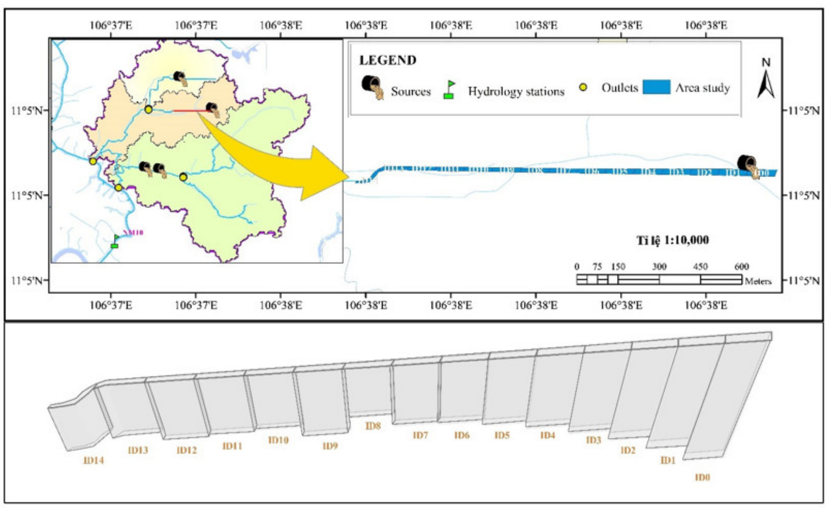

Figure 1: Study area

The concentrations of two parameters considered, BOD (mg/l) and DO, are given at cross-section ID0 as follows: BOD = 6.0 (mg/l), DO = 2.2 (mg/l), the temperature of the flow is equal to 15°C. The background concentration of the river is taken as follows: BOD = 6.0 (mg/l), DO = 2.2 (mg/l), and the river's background temperature is taken as 25°. Eddy coefficient is 0.25 (m^2/s) and Manning number coefficient is 34 ($m^{1/3}/s$). The problem is simulated based

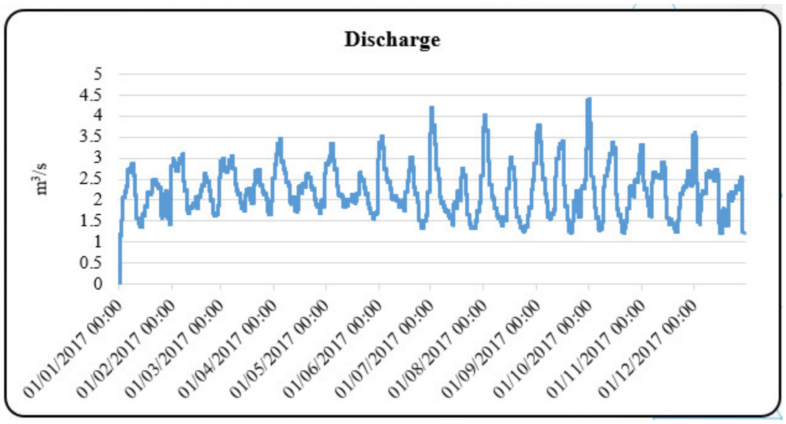

Figure 2: Discharge data

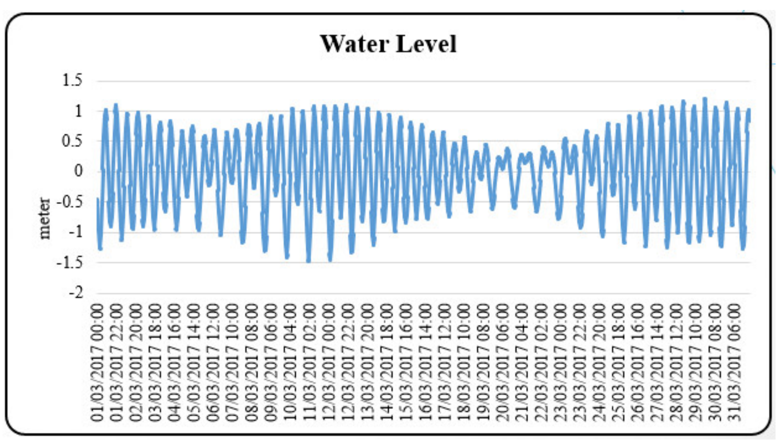

Figure 3: Water level boundary

on 2 scenarios, including: stable case scenario and incident case scenario. Stable: The discharger is located at location ID1 with a flow of 0.0266204 (m^3/s) (equivalent to 2300 m^3/day). The waste source temperature is taken as 20°C, the BOD concentration in the wastewater is 8.5 (mg/l), and DO = 1.5 (mg/l). For the incident scenario: there will be a change in 10 hours from 0:00 on March 1, 2017 to 9:00 on March 1, 2017 with the waste concentration values respectively: BOD = 20 mg/l, DO = 1 mg/l.

3.2.2 Extracting datasets for SPM

To ensure the dataset for SPM, the results of the flow water quality modelling by MIKE for two parameters BOD and DO were extracted from 0h00 am Feb 27, 2017 to 12h00 pm Mar 3, 2017 at location ID1, after which the values $L_0(t)$ and $D_0(t)$ as input data for SPM were calculated (Figure 4). The SPM running time is determined to be T=360 h (equivalent to 15 days); the SPM runtime step is taken to be t = 3600s; and the velocity u in the formula is the instantaneous velocity at each grid node obtained from the MIKE model. The results of calculating BOD and DO concentrations by MIKE were also extracted at the ID1 – ID13 nodes serving $L_1(x)$ and $D_1(x)$. Calculation results of BOD and DO concentrations were also extracted at 13 locations (ID1 – ID13), which were used to compare with the calculated results from SPM. The average running velocities at reach are extracted from MIKE's running results

3.3 Formula of k_2 and E

The first step in modelling a river network system is to divide the system into reaches, which are part of a stream with relatively uniform hydraulic properties. In this study, the selected river segment includes an integer number of reaches. To calculate the dispersion coefficient at the two-reach junction the following formula is used

$$E_{p;i} = 0.011\frac{U_i^2 B_i^2}{H_i U_i^*}$$

$$U_i^* = \sqrt{gH_iS_i}, g = 9.8m/s^2$$

where: $E_{p,i}$ is the longitudinal dispersion coefficient at the contiguous section between the two elements i and i+1 (m^2/s); U_i velocity (m/s) river reache i, B_i is reach width (m), H_i is the average depth of reach, S_i is reach slope. The reaeration coefficient is calculated based on the formula

$$k_2 = 10.9 \times \left(\frac{U}{H}\right)^{0.85}$$

Calculation results for the cases study shown in Table 1 and are used in the simulation of water quality using the Streeter – Phelps model in Section 3.1.

3.4 Error estimation

The notation for the solution $D(t,x)$, $L(t,x)$ obtained from solving the SPM is $D_{model}(x,t)$, $L_{model}(x,t)$. The tuples $D(t,x)$ and $L(t,x)$ obtained from the

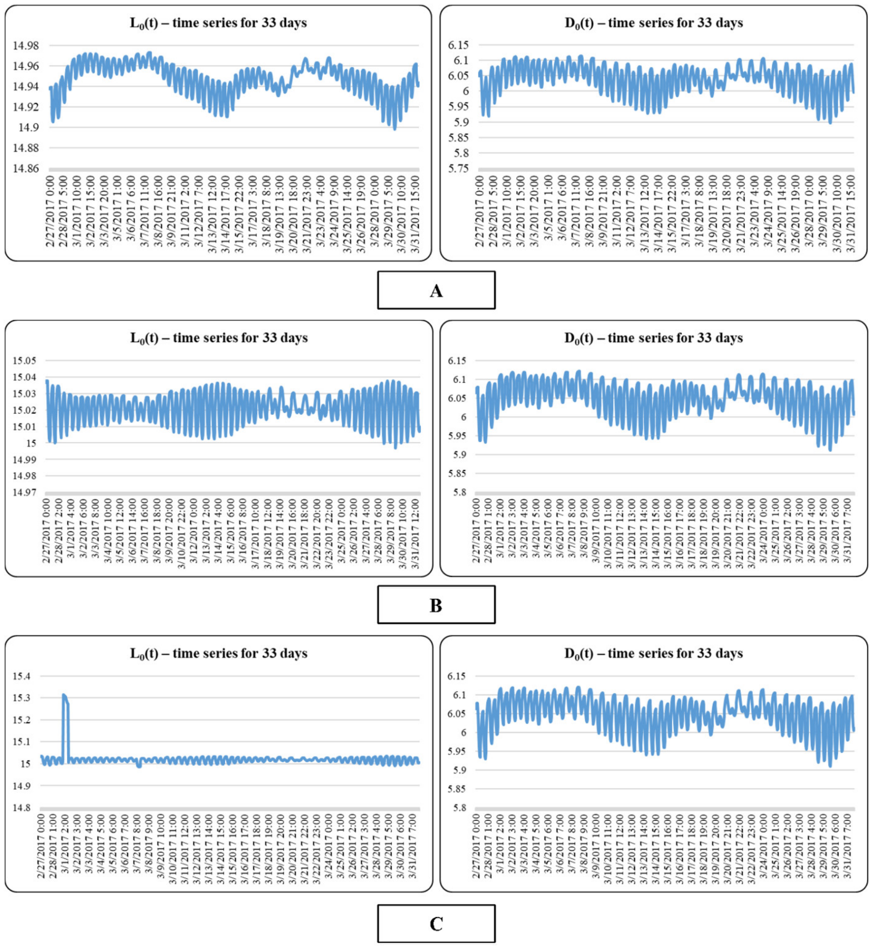

Figure 4: Dataset for MIKE EcoLab

MIKE model are D_{actual} and L_{actual}. The error between two datasets is determined by the distance calculated according to the Euclidean standard:

$$d\left(D_{model}, D_{actual}\right) = \left\| D_{model} - D_{actual} \right\|_2 = \sqrt{\sum_{i=1}^{M} \sum_{n=1}^{N} \left(D_i^n|_{model} - D_i^n|_{actual} \right)^2}$$

$$d\left(L_{model}, L_{actual}\right) = \left\| L_{model} - L_{actual} \right\|_2 = \sqrt{\sum_{i=1}^{M} \sum_{n=1}^{N} \left(L_i^n|_{model} - L_i^n|_{actual} \right)^2}$$

Position Node	E	k_2
ID1	0.0004182	0.169299
ID2	0.0006026	0.218226
ID3	0.0010324	0.292705
ID4	0.0017634	0.385719
ID5	0.0022595	0.498760
ID6	0.0029815	0.585598
ID7	0.0034534	0.650746
ID8	0.0070629	0.954641
ID9	0.0133378	1.331190
ID10	0.0139734	1.387406
ID11	0.0137954	1.434067
ID12	0.0105865	1.256748
ID13	0.0081569	1.386835

Table 1: Values of E and k_2

The error function is defined as:

$$\%D_{loss} = \frac{d\left(D_{model}, D_{actual}\right)}{\|D_{actual}\|_2}$$

$$\%L_{loss} = \frac{d\left(L_{model}, L_{actual}\right)}{\|L_{actual}\|_2}$$

3.5 Programming language

Step 1: Read data $L_{actual}(t_i, x_j)$, $D_{actual}(t_i, x_j)$ from excel file, where t_i is a time series separated by a time interval $\Delta t = 1h$, taking place over 15 days, from 0.00 am Feb. 27, 2017 to 12 pm Mar. 13, 2017, x_j with j running from 1 to 13; these are the cross-sectional positions from ID1 (where the waste source is) to ID14
Step 2: Take data of the initial conditions and boundary conditions to solve the system of Streeter-Phelps Equations
Step 3: Build a function to solve the system of equations to find the values of $L(t, x)$ and $D(t, x)$ at space-time steps: For i = 1 to M:

$$(L, D) = function(w) \text{ with } w = (\theta_1, \theta_2)$$

Step 4: Build the error function and output the error of (L_{model}, D_{model}) compared to (L_{actual}, D_{actual})
Step 5: Export data to the Excel file of L_{model} and D_{model}, and build a chart to compare SPM and MIKE results.

3.6 Framework and implementation steps

The first step is to apply MIKE to create the input dataset for SPM as well as to create a set of water quality simulation results to compare SPM and MIKE. The second step is related to SPM. Use MIKE outputs $D(t_0, ID_j)$ and $L(t_0, ID_j)$ to be the initial values for SPM. The third step, use the parameters k_1, k_2, E and the concentration of the waste source to run MIKE as well as SPM. Note that, for the SPM, k_2-values is at each position node. In the step 4, we evaluate the time series error $(L_{actual}(t_i, x_j), D_{actual}(t_i, x_j))$ and $(L_{Model}(t_i, x_j)$, $D_{Model}(t_i, x_j))$ (Figure 5). Then find the error value.

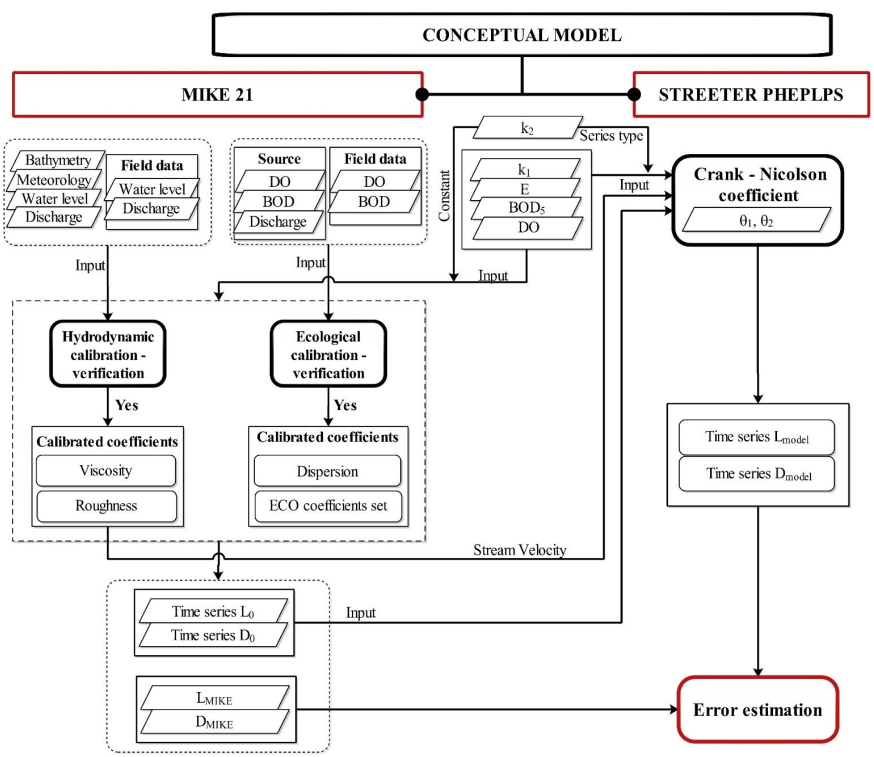

Figure 5: Study framework and implementation steps

4 Results and Discussion

4.1 Estimate the error between SPM and MIKE

The Crank–Nicolson algorithm presented in Section 3.1 is applied to solve systems (4) and (5). The space step is $\Delta x = 101$m; with time step: $\Delta t = 3600$s, L = 1212, $k_1 = 0.102$, calculated from ID1 to ID13; print 15 days. Input for SPM is given in the Table 1. Content and steps are taken according to Figure 5, detailed in Section 3.6. The error between the two models is shown in the last two columns, as well as in Figure 6. The error of L is relatively small in both scenarios, while the error of D is less than 6%.
The results of running SPM and evaluating the error of L_{model} and D_{model} compared to L_{actual} and D_{actual} are shown in the Table 2:

Case	k_1	θ_1	θ_2	$L_{loss}(\%)$	$D_{loss}(\%)$
0	0.102	0.5	0.5	0.5149	5.7697
1	0.102	0.5	0.5	1.0118	5.7764

Table 2: Results of running SPM for two cases

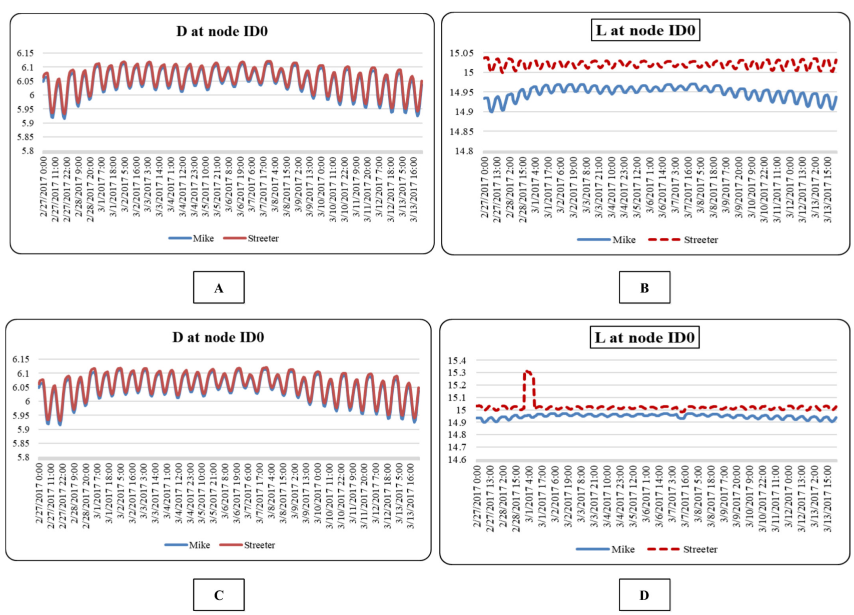

Figure 6: Compare the results of running MIKE and SPM at cross-sections ID0

4.2 Discussion

In the pollution control problem, parameter D plays an important role, which is the quantification of the pollution of the stream due to the wastewater discharger. The evaluation of error D allows a close relationship between the modelled and measured results. As a result, the error of D from the SPM solution depends on both the sampling distance and the sampling frequency.
The case 0 shows the relatively stable operation of the waste source, shown in the relatively stable $L_0(t)$ and $D_0(t)$. Case 1 shows the abnormality of the discharge source at some time, specifically on March 1, from 0.0 am to 9 am, the wastewater discharged into the river has a BOD_5 2.4 times higher than normal. The results of SPM and MIKE both show an increase in BOD and decrease in DO. The error for the variable L in case 0 reaches 0.5 is a relatively small error, in case 1, the error L reaches 1% can also be considered quite small. The error of variable D reaching $< 6\%$ is acceptable.

5 Conclusion

The forward problem in environmental modelling was studied. Based on the classical Streeter-Phelps equation, a problem of determining L and D representing wastewater dischargers and receptors for unsteady flow has been formed. The main results include the following:

- A numerical solution scheme for determination (L, D) has been given for the case of unsteady flows as well as waste dischargers based on the Crank – Nicolson scheme algorithm. The numerical solution was obtained. It is understood as the result of running the Streeter - Phelps model (SPM).

- Created datasets (L, D) for regular streams and wastewater dischargers. The MIKE model was used for a particular stream segment object. The dataset extracted from MIKE served as input for the SPM and was used to evaluate the error between SPM and MIKE.

- We performed numerical simulation (L, D) by using SPM according to different sets of parameters (k_1, k_2, E) and compared the results received with MIKE. The evaluation results show that SPM has good accuracy and can be used as an alternative to MIKE for unsteady flows.

- Performed experimental calculations with different spatial and temporal grid steps to clarify the influence of monitoring locations and measurement frequency. A number of rules have been identified, allowing us to shorten the number of measured samples as well as the sampling location.

The results of this study allow replacing the MIKE model with SPM in some cases, as well as serving a strong scientific basis for finding solutions to the inverse problem.

6 Acknowledgements

The authors acknowledge the support of time and facilities from Van Lang University (VLU), Viet Nam, for this study. The authors acknowledge the support of time and facilities from Ho Chi Minh City University of Technology (HCMUT), VNUHCM for this study.

References

[1] Streeter H. W. and B. Phelps, A study of the pollution and nature purification of the Ohio river, Factors concerned in the phenomena of oxidation and reaeration. Washington,D.C.: United states public health service, 1925.

[2] O'connor D. J., "Mechanism of Reaeration in Natural Streams," in Transactions of the American Society of Civil Engineers, Transactio., T. of the A. S. of C. Engineers, Ed. Transactions of the American Society of Civil Engineers: Transactions of the American Society of Civil Engineers, 1958.

[3] Haider H., W. Ali, and S. Haydar, "Evaluation of various relationships of reaeration rate coefficient for modeling dissolved oxygen in a river with extreme flow variations in Pakistan," Hydrol. Process., no. November 2008, pp. 2267–2274, 2012, doi: 10.1002/hyp.

[4] Omole D. O., E. O. Longe, and A. G. Musa, "An Approach to Reaeration Coefficient Modeling in Local Surface Water Quality Monitoring," Environ. Model. Assess., vol. 18, no. 1, pp. 85–94, 2013, doi: 10.1007/s10666-012-9328-0.

[5] Fu B. et al., Modeling Water Quality in Watersheds: From Here to the Next Generation, vol. 56, no. 11. 2020.

[6] Chapra S. C., "Engineering Water Quality Models and TMDLs," J. Water Resour. Plan. Manag., vol. 129, no. 4, pp. 247–256, 2003, doi: 10.1061/(asce)0733-9496(2003)129:4(247).

[7] USEPA, "Technical Guidance Manual for Developing Total Maximum Daily Loads," United States Environ. Prot. Agency, Off. Water, no. September, p. 254, 1995.

[8] USEPA, "Guidance for Water Quality-based Decisions: The TMDL Process," vol. EPA 440/4-, no. April, p. 62, 1991.

[9] Schaffner M., H. P. Bader, and R. Scheidegger, "Modeling the contribution of point sources and non-point sources to Thachin River water pollution," Sci. Total Environ., vol. 407, no. 17, pp. 4902–4915, 2009, doi: 10.1016/j.scitotenv.2009.05.007.

[10] 2017 DHI, Water Quality - ECO Lab WQ Templates. 2017.

[11] Tariq Ahmad Mahmoud and S. Ahmad, "A water quality study of a stretch of the River Tigris," Water Res., vol. 13, no. 8, pp. 785–790, 1979, doi: 10.1016/0043-1354(79)90243-4.

[12] Tyagi B., S. Gakkhar, and D. S. Bhargava, "Mathematical modelling of stream DO-BOD accounting for settleable BOD and periodically varying BOD source," Environ. Model. Softw., vol. 14, no. 5, pp. 461–471, 1999, doi: 10.1016/S1364-8152(98)00091-7.

[13] Pimpunchat B., W. L. Sweatman, G. C. Wake, W. Triampo, and A. Parshotam, "A mathematical model for pollution in a river and its remediation by aeration," Appl. Math. Lett., vol. 22, no. 3, pp. 304–308, 2009, doi: 10.1016/j.aml.2008.03.026.

[14] Fan F. M., A. S. Fleischmann, W. Collischonn, D. P. Ames, and D. Rigo, "Large-scale analytical water quality model coupled with GIS for simulation of point sourced pollutant discharges," Environ. Model. Softw., vol. 64, pp. 58–71, 2015, doi: 10.1016/j.envsoft.2014.11.012.

[15] Bui L. T., "Reaeration and Dispersion Coefficients Prediction for a River Flow by Implementing Machine Learning Algorithm on MIKE Dataset," 13th Int. Conf. Comput. Model. Simul. (ICCMS '21), 2021, doi: 10.1145/3474963.3474966.

[16] Avalos-Gauna E. and L. Palafox-Novack, "Heat transfer coefficient prediction of a porous material by implementing a machine learning model on a CFD data set," Int. Conf. Fluid Flow, Heat Mass Transf., pp. 1–10, 2019, doi: 10.11159/ffhmt19.149.

[17] López-Estrada F.-R. and G. Valencia-Palomo, "Advanced Mathematics and Computational Applications in Control Systems Engineering," in Mathematical and Computational Applications, vol. 26, no. 1, 2021, p. 20.

[18] Schnoor J. L., Environmental Modeling: Fate and Transport of Pollutants in Water, Air, and Soil. 1996.

[19] Tuan L. N., T. M. Quan, T. T. Thuy, D. T. Huy, and T. X. Hoang, "Assessing carrying capacity of receiving water bodies – A case study in Southern waterways of Binh Duong province," Sci. Technol. Dev. Journal, Nat. Sci., vol. 2, no. 6, pp. 84–97, 2018.

[20] DHI, MIKE 21 and MIKE 3 flow model FM, Hydrodynamic and Transport Module , Scientific Documentation. 2018.

[21] DHI, "Water quality - MIKE ECO Lab WQ Templates - Scientific Description," Hørsholm, Denmark, 2017.