Software for Pollutants Transport in Rivers and for Identification of Excessive Pollution Sources

Abstract

The program packages of realization of mathematical models of pollutants transport in rivers and for identification of river water excessive pollution sources located between two controlled cross-sections of the river will be considered and demonstrated. The software has been developed by the authors on the basis of mathematical models of pollutant transport in the rivers and statistical hypotheses testing methods. The identification algorithms were elaborated with the supposition that the pollution sources discharge different proportions of pollutants into the rivers. One-, two-, and three-dimensional advection-diffusion mathematical models of river water quality formation both under classical and new, original boundary conditions are realized in the package. New finite-difference schemes of calculation have been developed and the known ones have been improved for these mathematical models. At the same time, a number of important problems which provide practical realization, high accuracy and short time of obtaining the solution by computer have been solved. Classical and new constrained Bayesian methods of hypotheses testing for identification of river water excessive pollution sources are realized in the appropriate software. The packages are designed as up-to-date convenient, reliable tools for specialists of various areas of knowledge such as ecology, hydrology, building, agriculture, biology, ichthyology and so on. They allow us to calculate pollutant concentrations at any point of the river depending on the quantity and quality of river water excessive pollution sources when such necessity arises.

Keywords: Software; Pollutants transport; Identification; Excessive pollution sources; Mathematical models

Introduction

For solving the problems of analysis and control of the quality of the environmental objects, it is necessary to process big volume of measurement information about physical, chemical and biological parameters of these objects. To process big data with guaranteed quality in acceptable period of time is possible by means of wide application of mathematical methods and computers. For collecting, storage and processing data of the environment, there are developed the automated systems of the quality control of the environmental objects and universal software packages [1]. Among them the most widespread are environmental water and air pollution levels control systems. Among the most topical problems of control of environmental water quality, the task of modeling of polluting substances propagation in water objects should be emphasized.

Mathematical Models Describing the Pollutants Transport in Rivers

Transport and diffusion equation

The diffusion of polluting substances in rivers is mostly described by the three-dimensional equation of turbulent diffusion of non-conservative substances [1-4]:

\[
\frac{\partial \Phi(t, r)}{\partial t} - (\nabla \Phi(t, r)) + \phi(t, r) \Phi(t, r) + (r \nabla) \Phi(t, r) + \zeta(r) \Phi(t, r) = f(t, r),
\]

Where \( \Phi(t, r) \) is the time-averaged concentration of non-conservative pollutant; \( t \) is the time; \( r = [x, y, z] \) is the radius vector; its components \( x, y, z \) are the spatial coordinates (the axis \( x \) is horizontal and its direction coincides with the direction of averaged current of the flow, the axis \( y \) is perpendicular to the free surface and is directed downwards; the axis \( z \) is directed across the flow); \( K(r) \) is the tensor of turbulent diffusion, \( \zeta(r) \) is the vector of time-averaged speed of the river flow; \( f(t, r) \) is the coefficient of non-conservatism of pollutants, \( f(t, r) \) is the power of polluting sources.

For the engineering practice very often the two-dimensional or one-dimensional models are considered instead of the three-dimensional one [2,5]. If the non-uniformity of distribution of concentrations of pollutants on the depth of the watercourse is not taken into account, then it is possible to obtain the two-dimensional turbulent diffusion equation. The equation of one-dimensional turbulent diffusion is applied when the distribution of the concentration of the pollutant across the stream is homogeneous. It is also used when average indicators are used to represent pollution across the river [1].

*Corresponding author: Kachiashvili KJ, Georgian Technical University, Georgia
Kachiashvili KJ and Melikdzhanian DI
1Georgian Technical University, Georgia
2Tbilisi State University, Georgia

Received: August 26, 2016 | Published: November 16, 2016

MOJ Eco Environ Sci 2016, 1(1): 00006
Initial and boundary conditions

The unknown function $\Phi(t,r)$ is defined at $r \in G$ and $t \in T$, where $G$ is some region with the boundary $\partial G$, or $r \in T \subset [0,L]$ (in one-dimensional model).

The additional conditions are specified in the form of

$$\Phi(0,r)=\Phi_0, \quad \Phi(t,r)|_{r=\sigma}=\Phi(t,0)$$

($S_0, \sigma = \text{const}$). The boundary conditions in the lower extremity of the river section may be classical or non-classical. The classical conditions (condition of complete mixing) look like

$$\frac{\partial}{\partial \nu} \Phi(t,r)|_{r=L}=0; \quad \text{(condition of full mixing)}$$

non-classical conditions –

$$\Phi(t,r)|_{r=L}=q \Phi(t,r)|_{r=L-1}; \quad \text{(not local boundary condition)}$$

Here $q$ is the coefficient of self-cleaning of the river on the considered section, $L$ is the length of the section [5,6].

At using of two- or three-dimensional models, the following boundary conditions on the other part of the line or on the surface $\partial G$ (Neumann condition), are set also:

$$\left\{\begin{array}{l}
\nu \cdot \nabla \Phi(t,r) |_{r \in \partial G} = 0,
\end{array}\right.$$

where $\nu$ is unit vector of external normal to the border $\partial G$. In particular, at $m=3$, it should be

$$\frac{\partial}{\partial \nu} \Phi(t,r) |_{r=0} = 0.$$

Algorithms of solving of the initial-boundary-value problems

For solving the considered initial-boundary-value problems one of the following three algorithms is offered to the user at choice in the considered software [5]:

a. Using the classical explicit different scheme;

b. Using the classical implicit different scheme;

c. Using the method of decomposition of the operator.

The problems of optimum choosing of the algorithm parameters on which depend the accuracy, the time and the possibility of practical realization of the equation solution are realized in these algorithms [5,8].

The method of decomposition of the operator is the original algorithm of solving of the diffusion equation use in multidimensional problems. In this algorithm solution of multidimensional diffusion equation is represented in the form of a linear combination of solutions of some one-dimensional diffusion equations [5,7,8]. Thus, the multidimensional problem is reduced to the one-dimensional one. As the experience shows, the algorithm of decomposition of the operator can be realized by computer much more quickly than the classical explicit and implicit different schemes.

Description of river banks by splines

The problem of the analytical description of plane or spatial region for which the diffusion equations and the boundary conditions are investigated is solved [5,9]. This region represents the part of the channel of the river filled with water in the section for which we are modeling water pollution. The interpolation by splines is used for the analytical description of the plane curve, represented in the form of a sequence of distinct points with given Cartesian coordinates. Such a sequence, in particular, could be one of the river bank lines. The construction algorithms for splines of two types of the simplest explicit form which ensure continuous dependence of the tangent vector of the curve on its parameter are realized in the package. In particular, the river bottom is described by polynomial splines, and the river banks are described by trigonometrical splines [9,10].

Optimum choice of time and spatial steps of digitization

The problem of optimum choice of the steps of digitization of the algorithms of all realized different schemes is solved in the considered software [5,10,11]. The criteria of optimality are the requirements of reduction of the time and the errors of calculation as much as possible. The algorithms are proper for the functions describing the polluting substances transport in water having the derivatives up to the fourth order inclusive.

The total number of nodal points $N$ in the difference schemes determines the time necessary for realization of the algorithm; the accuracy of the obtained result depends on it. Naturally, there arises the question: how to select the numbers $n_l, \ldots, n_m$ at the given value of $N$ so that the algorithm should be somewhat optimum.

This problem is solved in the following way [11]: at the given values of total nodal points $N$, there are determined such values of spatial steps of the grid along coordinate axes $h_k \rightarrow h_p$ for which the upper bound of the module of the residual takes on the minimum value:

$$\sum_{k=1}^{n} \Theta_k h_k^2 \rightarrow \min,$$

$$\prod_{k=1}^{m} h_k^2 = H = \text{const},$$

Where $\Theta$ is the error of approximation of the equation along the $k$-th coordinate ($k=1,\ldots,m$).

The solution of this optimization problem is determined by the formula:

$$h_k^2 = \frac{1}{\Theta_k} \left( \prod_{k=1}^{m} \Theta_k h_k^2 \right)^{1/m} (k=1,\ldots,m).$$

The problem of optimum choice of the value of the parameter $\tau$ at given spatial steps of the grid is solved in the following way:
at the given value of upper bound of the module of the residual of the equation which is equal to \( \tau \), there are determined such values of \( N \) and \( N' \) for which the time of calculation takes on the minimum value:

\[
N^2 ; t \rightarrow \min ,
\]

\[
c_t \tau^2 + PN^{-2/w} = \varepsilon = \text{const} .
\]

Where \( \Theta \) is the error of approximation of the equation by the time coordinate; \( K \) is the constant dependent on the algorithm of solution of the sets of linear equations, usually \( k = 1 \); \( \varepsilon \) is the given value of the upper bound of the module of the equation residual.

The solution of this optimization problem is determined by the formula:

\[
N = \frac{P(k+1/m)}{k/\varepsilon} ; \quad \tau^2 = \frac{1}{c_t} 1 + km ;
\]

\[
c_t k^2 = \ldots = c_t k_m = k c, \tau^2 = \frac{k c}{1 + km} .
\]

The main difficulty for practical realization of the described scheme of optimization is connected with the necessity in estimation of the upper bound of partial derivatives of the function \( \Phi(t, r) \) with respect to the spatial coordinates in terms of which parameters \( \Theta \) and \( \Theta_b \) are expressed without solving the diffusion equation.

**Estimation of derivatives of the unknown function**

For practical realization of the described optimization schemes, it is necessary to estimate somehow the upper bounds of the modules of partial derivatives of the unknown function with respect to independent variables without solving the initial task. It should be taken into account that the derivatives of these functions can be estimated with the accuracy of the common constant multiplier.

One way of solution of this problem consists in the following: for some values of the parameters \( h_k \) and \( \Phi \) (at solving the diffusion equation), the values of the function \( \Phi \) are determined as the first approximation. Then, using the numerical differentiation operations, the required derivatives are determined and the values of the desired parameters are calculated. Using these values, new values of the function \( \Phi \) are determined and so on until the difference between the neighboring calculated values of the function are less than the given value [12].

At solving the diffusion equation, it is possible to use the explicit scheme at the first stage. Then the determination of the values of the sought for function with higher accuracy is necessary. In this case, for calculation of the unknown function values, the iteration method similar to the one used in work [1] for calculation of the multidimensional integral by the Monte-Carlo method is applied. It is known that the operations of numerical differentiation are not stable, but this should not prevent the realization of the offered method since it requires only rough estimates of unknown derivatives. An obvious drawback of this method is the necessity in performance of a plenty of additional actions. Though, due to the capabilities of modern personal computers, it is negligible at solving the particular tasks.

**The effect of smoothness of the inhomogeneous part of the diffusion equation on the accuracy of the results**

The inhomogeneous parts of diffusion equation (1) realized in the package contain the impulse functions, which are linear combinations of Dirac delta-functions. These functions and their derivatives are not bounded. Therefore the difference schemes of the solution of diffusion equations are not correct in the vicinities of pollution source localization points.

For elimination of this drawback, it is necessary to consider a model, in which the delta function \( \delta(x-a) \) is replaced by the bounded numerical function \( D(u, x-a) \), where \( u \) is the additional parameter. It means that the point sources are replaced by extended ones, the capacity of each of which has the maximum corresponding to the capacity of the source at the point of its location. Such sources can be named quasi-point sources.

The function \( D(u, x) \) should satisfy the following requirements: it reaches the maximum value proportional to \( 1/u \) at the point \( x=0 \); it tends to zero at \( x \rightarrow \pm \infty \); its integral between the limits \( \pm \infty \) and \( +\infty \) is equal to unit; its plot represents a bell-shaped curve. The width of such curve is naturally defined as the width of the rectangle which height is equal to the ordinate of the peak of the curve, and its area is equal to the area of the figure limited by the curve and the axis of abscissas. Thus, the parameter \( u \) characterizes the width of the plot of the function \( D(u, x) \), i.e. the sizes of the source. At \( u \rightarrow 0 \), the function \( D(u, x) \) tends to \( \delta(x) \).

One of the elementary continuous functional dependences which may be used for setting the function \( D(u, x) \) is the trapezoid dependence:

\[
D(u, x) = \begin{cases} 1/u & \text{at } t < 1-s, \\ (1-s-t)/(2us) & \text{at } 1-s \leq t \leq 1+s, \\ 0 & \text{at } t > 1+s, \end{cases}
\]

Where \( t = 2s)/u ; s \) is any real parameter from the interval \((0,1)\). The plot of this function together with the axis of abscissas forms an isosceles trapezium, the top and bottom bases of which have the lengths equal to \( u(1-s) \) and \( u(1+s) \), respectively, and the height of which is equal to \( 1/u \). The less is value \( s \), the less this trapezium differs from a rectangle.

Besides the trapezoid dependence, in the developed software package the following dependences are offered to the user’s choice:

- **Gaussian**

\[
D(u, x) = \frac{2}{2\pi u(1+(x/u)^2)} ;
\]
b. Lorentzian

\[ D(u, x) = \frac{2}{2\pi u(1+(x/u)^2)} . \]

In fact, the function \( D(u, x) \) defined by one of the two latter relations plays the part of a smoothing function, which allows us to transform any function from \( L^2 \) into an infinitely differentiable function by means of the operation of convolution [10,11].

If the pollution sources do not operate continuously, but they operate for a limited interval of time \([0,T]\), then the capacity of each source should be smoothed not only by spatial coordinates, but also by time. It means that the function should be proportional to \( A(v,t-T) \), where \( v \) is the additional parameter, and \( A(v,t) \) is the function satisfying the following conditions: at \( t \rightarrow \infty \) it tends to unit, and at \( t \rightarrow -\infty \) to zero. Its plot represents a quasi-stepped curve, which steepness is maximum at \( 1/v \), and this maximum value of the steepness is proportional to \( 1/v \). The parameter \( v \) characterizes the time of diminution of the function of discharge. At \( A(v,x) \), the function \( A(v,x) \) tends to \( 1-\delta(x) \), where \( \delta(x) \) is Heaviside's stepped function. In the developed software package, it is possible to choose an explicit form of the function \( A(v,x) \) from the following types:

a. 

\[ A(v,t)=\frac{1}{2\pi} \frac{1}{\arctan(t/v)} . \]

As the computation results show, the smoother are the functions of discharge, the more accurate are the results of solution of the equations in the vicinities of the points of discharge. For given types of functions \( D(u,x) \) and \( A(v,t) \), the accuracy of the results increases as the values of parameters \( u \) and \( v \) increase.

Capabilities of the package

Input and editing of the initial data

The input of information necessary for operation of the package is performed in the form of separate constants, tables or the choice of the appropriate line from the pull-down. Data for editing are chosen by means of teams of the main menu (Figure 1).

The entered data which form sequences of elements of the same type are represented in the form of tables (Figure 2). When there are several editing windows with tables on the screen, different windows correspond to different sections of the river.

Realization of computation and representation of the results

The results of calculations are displayed in the form of text messages, tables, graphs, etc. For example, the formats of display of computation results as diagrams and tables are shown in Figure 3 & 4, respectively.

The represented text and diagrams may be printed or written in a text- or graphic files, the name of which is determined by default or indicated by the user.

![Figure 1: Data for editing are chosen by means of teams of the main menu.](image)
Figure 2: The entered data which form sequences of elements of the same type are represented.

Figure 3: The results of calculations are displayed in the form of text messages.
Software for Identification of Excessive Pollution Sources

The software for identification of river water excessive pollution sources located between two controlled cross-sections of the river has been developed by the authors on the basis of mathematical models of pollutant transport in the rivers (described above) and statistical hypotheses testing methods. Depending on the available a priori information, different methods of statistical hypotheses testing can be chosen in the software for identification. With increasing of a priori information, application of more complicated methods becomes possible, thus providing their higher assurance. The following statistical tests are realized in this package:

A. The Euclidean distance method;
B. The Mahalanobis distance method;
C. The unconstrained Bayesian method with the stepwise and arbitrary loss functions;
D. The Constrained Bayesian Method (CBM);
E. The quasi-optimal test based on CBM.

The peculiarities of CBM methods are described in [13-18]. Therefore we will not spend space here for their description. Other methods are widely known and described in many scientific works. Let’s show only some characteristics of the package. In particular, in Figure 5 are shown the modes realized in the package, the input of the initial data and the calculation regimes. In Figure 6 is shown the graphical view of package output information where the 3-rd point in the bed of the river corresponds to the pollution source guilty of the excessive pollution (Figure 6).

Discussion

The results of detailed experimental research of algorithms and appropriate programs included in computer packages described above are given in [5]. There are considered the results of research of algorithms of calculation of concentration of polluting substances in the rivers with the help of diffusion equations and also sensitivity of models of different dimensions to the geometrical sizes of a section of the river and dependence of quality of identification of emergency pollution sources on the level of the noise, deforming results of measurement. They show truth and reliability both the created algorithms and programs of their realization [19-23].
From the results of calculations, the following practical recommendation concerning the use of the developed models may be given. If geometric sizes, locations of pollution sources, hydrologic characteristics and pollution conditions of the river are such that full mixing of water takes place upstream of the lower controlled section, then it is enough to use the one-dimensional river models, which are considerably faster and require less computer memory than the models of greater dimensionality. Otherwise it is necessary to use the models of greater dimensionality. When choosing a working model for a concrete section of a certain river, it is necessary to perform preliminary studies with due regard for the above factors. The proper choice of the model and its parameters ensures the qualitative modeling and identification of excessive discharge sources.

**Conclusion**

The computer packages of realization of mathematical models of propagation of polluting substances in rivers and for identification of excessive pollution sources of the river have been created. The packages comprise original models, methods and algorithms developed by the authors. The software packages are realized for IBM-compatible personal computers according...
to the generally accepted standard for similar production all over the world. The consumer can use them as a modern, convenient, simple and reliable tool for resolution the problems he deals with in the considered area. Versatile experimental investigation of the developed software packages and the algorithms realized in them, have confirmed their high computing, operational and service qualities. They are good tools for the experts of different professions for qualified solving many practical problems.

References

1. Primak AV, Kafarov VV, Kachiashvili KJ (1991) System analysis of air and water quality control. Naukova Dumka: Kiev, Ukraine.
2. Karashev AV (1969) River hydraulic. Leningrad: Hydrometeoizdat, Canada.
3. Nikanorov AM, Trunov NM (1999) Intrareservoir processes and quality control of natural water. Hydrometeoizdat: St.-Petersburg, Canada.
4. Gallakher L, Khobbs DD (1981) Dissemination of pollutants in estuary: mathematical models of the control of water pollution. Moscow: Mir p. 220-243.
5. Kachiashvili KJ, Melikdzhanian DI (2012) Advanced Modeling and Computer Technologies for Fluvial Water Quality Research and Control. Nova Science Publishers, New York.
6. Gordeziani DG, Gordeziani EG, Kachiashvili KJ (1999) About some mathematical models of propagation of solutions in rivers. IX intergovernmental conference Problems of an ecology and exploitation of objects of power engineering. Sevastopol, p. 77-79.
7. Kachiashvili KJ, Gordeziani, DG, and Melikdzhanian DI (2001) Mathematical models of Pollutants Transport with Allowance for Many Affecting Pollution Sources. Urban Drainage Modeling Symposium, Florida, pp. 692-702.
8. Gordeziani DG, Samarskii AA, (1978) Some problems of the thermo-elasticity of plates and shells, and the method of summary approximation. Complex analysis and its applications, Moscow, pp. 173-186.
9. Kachiashvili KJ, Melikdzhanian DI (2002) Analytical description of the bank line of the river for simplification and improvement of the process of calculation of polluting substances concentration. In: Reports of enlarged session of the workshop of I. Vekua Institute of Applied Mathematics 17(3): 101-109.
10. Kachiashvili KJ, Gordeziani DG, Lazarny RG, Melikdzhanian DI (2007) Modeling and simulation of pollutants transport in rivers. Applied Mathematical Modelling 31(7): 1371-1396.
11. Kachiashvili KJ, Melikdzhanian DI (2006) Parameter optimization algorithms of difference calculation schemes for improving the solution accuracy of diffusion equations describing the pollutants transport in rivers. Applied Mathematics and Computation 183(2): 787-803.
12. Kachiashvili KJ, Melikdzhanian DI, Prangishvili, Al (2015) Computing Algorithms for Solutions of Problems in Applied Mathematics and Their Standard Program Realization: Part 1-Deterministic Mathematics. Nova Science Publishers, New York, USA.
13. Kachiashvili KJ, Mued A (2013) Conditional Bayesian Task of Testing Many Hypotheses. Statistics 47(2): 274-293.
14. Kachiashvili KJ (2011) Investigation and Computation of Unconditional and Conditional Bayesian Problems of Hypothesis Testing. ARPN Journal of Systems and Software 1(2): 47-59.
15. Kachiashvili GK, Kachiashvili KJ, Mued A (2012) Specific Features of Regions of Acceptance of Hypotheses in Conditional Bayesian Problems of Statistical Hypotheses Testing. Sankhya A: The Indian Journal of Statistics 74(1): 112-125.
16. Kachiashvili KJ, Hashmi MA, Mued A (2012) Sensitivity Analysis of Classical and Conditional Bayesian Problems of Many Hypotheses Testing. Communications in Statistics -Theory and Methods 41(4): 591-605.
17. Kachiashvili KJ, Hashmi MA (2010) About Using Sequential Analysis Approach for Testing Many Hypotheses. Bulletin of the Georgian Academy of Sciences 4(2): 20-25.
18. Kachiashvili KJ, Hashmi MA, Mued A (2013) Quasi-optimal Bayesian procedures of many hypotheses testing. Journal of Applied Statistics 40(1): 103-122.
19. Kachiashvili KJ (2014) Comparison of Some Methods of Testing Statistical Hypotheses: Part I. Parallel Methods. International Journal of Statistics in Medical Research 3(2): 174-189.
20. Kachiashvili KJ (2014) Comparison of Some Methods of Testing Statistical Hypotheses: Part II. Sequential Methods. International Journal of Statistics in Medical Research 3: 189-197.
21. Kachiashvili KJ (2014) The Methods of Sequential Analysis of Bayesian Type for the Multiple Testing Problem. Sequential Analysis 33(1): 23-38.
22. Kachiashvili KJ (2014) Investigation of the method of sequential analysis of Bayesian type. Journal of Advances in Mathematics and Computer Science 47(4): 1367-1380.
23. Kachiashvili KJ (2015) Constrained Bayesian Method for Testing Multiple Hypotheses in Sequential Experiments. Sequential Analysis: Design Methods and Applications 34(2): 171-186.