

Fuzzy Waste Load Allocation Model: Simulation-Optimization Approach

Pradeep P. Mujumdar¹ and V. R. Subbarao Vemula²

Abstract: The problem of waste load allocation (WLA) for water quality management of a river system is addressed with a simulation-optimization approach. The WLA model developed in the study provides the best compromise solutions to the pollution control agency (PCA) responsible for maintaining the water quality and the dischargers disposing pollutants into the river system. A previously developed fuzzy waste load allocation model (FWLAM) is extended to incorporate QUAL2E, a water quality simulation model developed by the U.S. Environmental Protection Agency for modeling the pollutant transport in a river. The imprecision associated with establishing water quality standards and the aspirations of the PCA and dischargers are quantified using fuzzy goals with appropriate membership functions. The membership functions of the fuzzy goals represent the variation of the goal satisfaction in the system. A genetic algorithm (GA) is used as an optimization tool to find optimal fraction removal levels to the dischargers and the corresponding satisfaction level. Because a GA is an unconstrained optimization tool, it is extended to handle constraints by complementing it with homomorphous mapping (HM), a constraint handling method for evolutionary algorithms. The GA directs the decision vector in an encoded form to HM. HM, after a few interactions with QUAL2E, redirects the decoded solution back to the GA. The GA assigns a fitness value to the feasible solution vector and applies operators to refine the solution. This interaction among the GA, HM, and QUAL2E continues until a prespecified criterion for global optimality is met. Application of the model is illustrated with a case study of the Tunga-Bhadra River in South India.

DOI: 10.1061/(ASCE)0887-3801(2004)18:2(120)

CE Database subject headings: Water quality; Fuzzy sets; Evolutionary computation; Optimization; River systems; Simulation.

Introduction

Waste load allocation (WLA) refers to determining the required pollutant treatment levels at a number of point and nonpoint sources to attain a satisfactory water quality response in a receiving water body in an economically efficient manner. A WLA model is, in general, a mathematical model incorporating a water quality simulation model within the framework of multiobjective optimization. It normally consists of three components: (1) an optimization model representing the objectives, goals, and constraints of the water quality management problem; (2) a water quality simulation model that yields spatial distribution of the water quality constituents in the river system; and (3) means of addressing uncertainty inherent in the system. Generally, two sets of objectives are considered in the decision-making process for water quality management of a river system. The first set of objectives deals with satisfying water quality standards set by the pollution control agency (PCA). The second set of objectives deals with the minimization of waste treatment cost to be borne by the dischargers in the river system. These two sets of objectives are often in conflict with each other.

Most WLA models include the Streeter-Phelps (S-P) equations (Streeter and Phelps 1925) with Camp-Dobbins (Camp 1963; Dobbins 1964) modifications to simulate biodegradation and to map waste loads into downstream dissolved oxygen (DO) concentrations. While S-P equations are effective in modeling DO and biochemical oxygen demand (BOD), they cannot be extended to model transport of other constituents (e.g., nitrogen, phosphorus, and chlorophyll). Several simulation models are now available (e.g., QUAL2E and WASP4) for modeling transport of most pollutants in a river system. Efforts to incorporate such simulation models in WLA models began with Cardwell and Ellis (1993), who addressed model uncertainty considering different models [S-P equations, QUAL2E (Brown and Barnwell 1987), and WASP4 (Robert et al. 1988)] simultaneously in a single framework.

Optimization methods have been developed to incorporate multiple and conflicting goals of the dischargers and the PCAs in the WLA models. Recently, Suresh and Mujumdar (1999), Sasi-kumar and Mujumdar (2000), and Mujumdar and Sasikumar (2002) have incorporated multiple and conflicting goals in the WLA models and addressed uncertainty due to both randomness and imprecision. Their models, however, lack realistic water quality simulation models, necessary to become complete enough to address all major issues of WLA problems. In this paper, a simulation-optimization (S-O) approach is proposed to integrate the fuzzy waste load allocation model with a water quality simulation model. The proposed methodology is different from the S-O method proposed by Burn (1989), Burn and Lence (1992), and Takyi and Lence (1994, 1999), in which simulation essentially meant a Monte Carlo simulation.

Several advantages of the S-O methodology have been realized in various fields of water resources, including groundwater

¹Associate Professor, Dept. of Civil Engineering, Indian Institute of Science, India. E-mail: pradeep@civil.iisc.ernet.in

²Scientist, Aeronautical Development Establishment, Defence Research and Development Organization, Bangalore, India. E-mail: vvrsubbarao@yahoo.com

Note. Discussion open until September 1, 2004. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on April 26, 2002; approved on May 12, 2003. This paper is part of the *Journal of Computing in Civil Engineering*, Vol. 18, No. 2, April 1, 2004. ©ASCE, ISSN 0887-3801/2004/2-120-131/\$18.00.

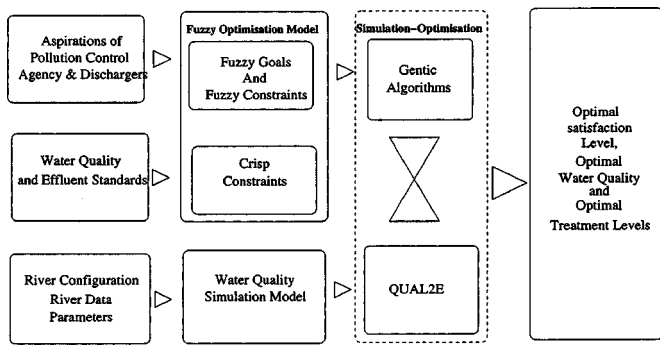


Fig. 1. Organization of fuzzy waste load allocation model

management (Gorelick et al. 1984; McKinney and Lin 1994), surface water quality and quantity management (Dai and Labadie 2001), water distribution systems (Sakarya and Mays 2000), and reservoir operation (Oliveira and Loucks 1997). A major advantage of the S-O methodology proposed in this paper is that the physical processes such as the mass and temperature balance are accounted through simulation outside the optimization model, thus reducing the size and complexity of the optimization model. The S-O method proposed in this paper is similar to the model suggested by Carmichael and Strzepek (2000). They linked QUAL2E with the MINOS package (Murtagh and Saunders 1980, 1982; Gill et al. 1981), which was used as an optimization tool. For a given pollution abatement matrix, the water quality model, QUAL2E, calculates the Jacobian matrix whose elements represent the marginal effects of increase in each pollutant load on downstream DO levels of the river. The overall optimization model then solves for the least cost matrix from the values of Jacobian matrix and provides an improved solution matrix to the simulation model.

The model presented in this paper incorporates simple and readily available packages [QUAL2E, PGAPack (David 1996), and GENOCOP (Koziel and Michalewicz 1999)] to reduce considerably the effort in development of the WLA model. The main objective of the work presented in this paper is to demonstrate applicability of an S-O methodology that incorporates a recent optimization tool, a water quality simulation model, multiple and conflicting objectives, and uncertainty due to imprecision in a single modeling framework. An outline of the model is shown in Fig. 1. The fuzzy optimization model takes into account the goals of the PCA and dischargers, stated subjectively and in an imprecise manner. The physical processes of the river system are modeled using the water quality simulation model, which yields water quality levels at specified monitoring points, or checkpoints, for known pollutant loadings, reaction rate coefficients, and travel time of a pollutant from its source to the checkpoint. Historically, Camp-Dobbins modifications to the S-P equations have been used to model BOD and DO transport in streams. The S-P equations, being linear formulations, permit use of linear programming (LP) as an optimization tool. Linear approximations may, however, lead to a simplified BOD-DO model and may result in high model uncertainty. The present work incorporates QUAL2E, a water quality simulation model to generate the river system response to a given set of pollutant loadings. QUAL2E, being a separate external simulation model, precludes the use of LP as an optimization tool. Also, it is well known that classical, nonlinear optimization methods pose a difficulty in achieving global or near global optimal solutions. As an alternative, therefore, the genetic algorithm (GA), an evolutionary algorithmic technique that assures

global or near global solutions, has been employed to solve the fuzzy optimization problem. QUAL2E and the GA are used in a simulation-optimization framework. Because GA is an unconstrained optimization method, homomorphous mapping (HM) (Koziel and Michalewicz 1999), a constraint handling technique of evolutionary algorithms that assures a feasible solution, if one exists, is used. The combined S-O model is developed in a deterministic framework (i.e., for a fixed flow and quality characteristics of river and effluent flows), with uncertainty due to imprecision addressed by treating the goals of the PCA and dischargers as fuzzy. The following section gives a description of a WLA problem and provides an overview of the fuzzy waste load allocation model (FWLAM) developed by Sasikumar and Mujumdar (1998), which is used in the present work.

Fuzzy Waste Load Allocation Model

The FWLAM is described using a general river system. The river consists of a set of dischargers who are allowed to release pollutants into the river after removing some fraction of their constituents. These fraction removal levels are necessary in order to maintain the acceptable water quality condition in the river as prescribed by the PCA. The acceptable water quality condition is ensured by checking water quality indicator levels (e.g., DO concentration) at a finite number of locations, which are referred to as checkpoints. The water quality indicator is affected by the controllable, as well as the uncontrollable, source of pollutants in the river system. In a water quality management problem, the concentration of the water quality indicator is expressed in terms of the fraction removal levels for the pollutants released by the dischargers in the river system. Considering the goals of the dischargers and the PCA as fuzzy goals, an optimization model is formulated with the set of fraction removal levels forming the decision variables. The resulting fuzzy optimization problem is stated as follows:

$$\text{Maximize } \lambda \quad (1)$$

subject to

$$\mu_{E_{il}}(c_{il}) \geq \lambda \quad \forall i, l \quad (2)$$

$$\mu_{F_{imn}}(x_{imn}) \geq \lambda \quad \forall i, m, n \quad (3)$$

$$c_{il}^L \leq c_{il} \leq c_{il}^D \quad \forall i, l \quad (4)$$

$$\max[x_{imn}^L, x_{imn}^{\text{MIN}}] \leq x_{imn} \leq \min[x_{imn}^M, x_{imn}^{\text{MAX}}] \quad \forall i, m, n \quad (5)$$

$$0 \leq \lambda \leq 1 \quad (6)$$

where λ = minimum satisfaction level in the system; c_{il} = concentration of water quality indicator i at checkpoint l of the river system; $\mu_{E_{il}}(c_{il})$ = fuzzy membership function of the goal of the PCA with respect to the water quality level c_{il} ; x_{imn} = fraction removal level at discharger m for pollutant n for control of water quality indicator i ; $\mu_{F_{imn}}(x_{imn})$ = fuzzy membership function of the goal of the discharger m with respect to the fraction removal level x_{imn} ; c_{il}^D = desirable level; and c_{il}^L = minimum permissible level for water quality indicator i at checkpoint l ($c_{il}^D \geq c_{il}^L$). The aspiration (i.e., minimum desired) level and maximum fraction removal level acceptable to the discharger m with respect to x_{imn} are represented as x_{imn}^L and x_{imn}^M , respectively. The PCA imposes minimum fraction removal levels that are also expressed as the lower bounds x_{imn}^{MIN} in constraint Eq. (5). The upper bound x_{imn}^{MAX} of the same constraint represents the

technologically feasible maximum fraction removal level. Observing that the maximum acceptable level of pollutant treatment cannot exceed the technologically possible upper limit, x_{imn}^M is always considered the upper bound of the constraint. The fuzzy goal of the PCA is to make the concentration level c_{il} of water quality indicator i at checkpoint l as close as possible to the desirable level, c_{il}^D , so that the water quality at checkpoint l is enhanced with respect to the water quality indicator i , for all i and l . If the DO level is the water quality indicator, a nondecreasing membership function suitably represents the fuzzy goal of the PCA. Similarly, the goal of the discharger is to make the fraction removal level x_{imn} as close as possible to the aspiration level x_{imn}^L for all i, m , and n . A nonincreasing membership function is suitable to reflect the goal of the discharger. The objective function of the FWLAM is interpreted as a maximization of the minimum satisfaction level in the system.

Substituting expressions for the membership functions of the fuzzy goals $\mu_{E_{il}}(\cdot)$ and $\mu_{F_{imn}}(\cdot)$ into constraint Eqs. (2) and (3), respectively, the mathematical formulation of the fuzzy optimization problem is expressed as follows:

$$\text{Maximize } \lambda \quad (7)$$

subject to

$$[(c_{il} - c_{il}^L)/(c_{il}^D - c_{il}^L)]^{\alpha_{il}} \geq \lambda \quad \forall i, l \quad (8)$$

$$[(x_{imn}^M - x_{imn})/(x_{imn}^M - x_{imn}^L)]^{\beta_{imn}} \geq \lambda \quad \forall i, m, n \quad (9)$$

$$c_{il}^L \leq c_{il} \leq c_{il}^D \quad \forall i, l \quad (10)$$

$$\max[x_{imn}^L, x_{imn}^{\text{MIN}}] \leq x_{imn} \leq x_{imn}^M \quad \forall i, m, n \quad (11)$$

$$0 \leq \lambda \leq 1 \quad (12)$$

The exponents α_{il} and β_{imn} are nonzero, positive, real numbers. Assignment of numerical values to these exponents is subject to the desired shape of the membership functions and may be chosen appropriately by the decision maker. The concentration level, c_{il} , of water quality indicator i at checkpoint l can be related to the fraction removal level, x_{imn} , of pollutant n from discharger m to control the water quality indicator i . This may be mathematically expressed as

$$c_{il} = \sum_{m=1}^{M_u} \sum_{n=1}^{N_u} f_{ilmn}(L_{ilmn}, x_{imn}) \quad (13)$$

where L_{ilmn} = concentration of pollutant n prior to treatment from discharger m that affects the water quality indicator i at checkpoint l . M_u and N_u are, respectively, the total number of dischargers and pollutants upstream of checkpoint l that influence c_{il} . The transfer function $f_{ilmn}(\cdot, \cdot)$ of Eq. (13) indicates the aggregate effect of all pollutants and all dischargers located upstream of checkpoint l on water quality indicator i . The transfer function can be evaluated using appropriate mathematical models that determine spatial distribution of the water quality indicator due to pollutant discharge into the river system from point sources. Depending on the nature of the water quality simulation model used, the transfer function may be either linear or nonlinear functions of the fraction removal levels. In the present work, the water quality simulation model QUAL2E, developed by the U.S. Environmental Protection Agency (EPA), is used for obtaining the concentrations c_{il} for given values of the fraction removal levels, x_{imn} .

Water Quality Simulation Model

A water quality model is a set of mathematical expressions defining the physical, biological, and chemical processes that take

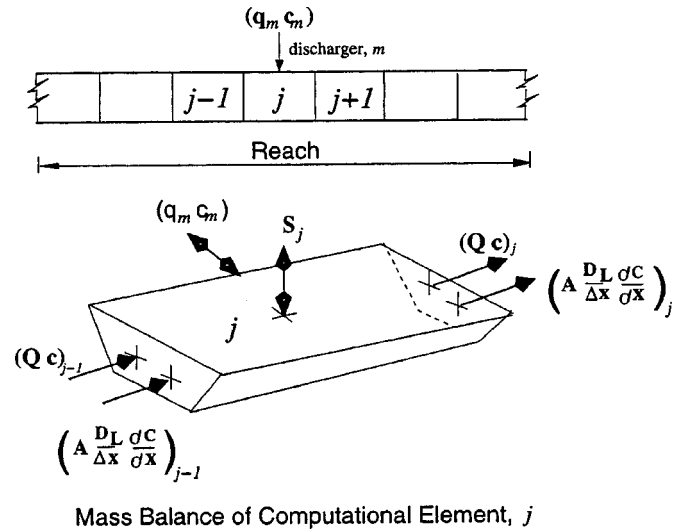


Fig. 2. River system discretization

place in a water body. Most river water quality models are developed based on the conservation of mass. Complex multiconstituent water quality models have been proposed and applied to predict the physical, chemical, and biological interactions of many constituents and organisms found in natural water bodies (e.g., Canale 1976; Harleman 1978). QUAL2E and WASP4 are two such models that are widely used for screening various policies in water quality management problems. In the present study, QUAL2E is used as a simulation model to obtain the water quality concentrations. It may be noted that the method proposed in this paper is general in its applicability and can be extended to include any other water quality simulation model.

QUAL2E has been used as a water quality simulation model in the recent works of Cardwell and Ellis (1993), De Azevedo et al. (2000), Burn and Yulianti (2001), Han et al. (2001) and Maier et al. (2001). In QUAL2E, a river is discretized into small and uniform elements, j ; a mass balance equation is applied to compute the concentration of a water quality indicator. The mass transformations across an element j is shown in Fig. 2 and is expressed for a particular water quality constituent as

$$\frac{\partial M_j}{\partial t} = \left[\left(A D_L \frac{\partial c}{\partial x} \right)_j - \left(A D_L \frac{\partial c}{\partial x} \right)_{j-1} \right] - [(A \bar{u} c)_j - (A \bar{u} c)_{j-1}] + V_j \frac{dc_j}{dt} + S_j \quad (14)$$

where M_j = mass of the constituent of concentration c_j in element j . The length, cross-sectional area, and volume of the element are Δx_j , A_j , and v_j , respectively. The dispersion coefficient of flow in the element is represented by D_j . u_j and Q_{j-1} indicate the mean velocity of flow after entering the element with mass flow rate Q_{j-1} . s_j is the external source and sinks of the constituent mass. Eq. (14) states that the accumulation of mass over time, $\partial M_j / \partial t$, in element j is equal to the difference between the flux entering the element, $[A(D_L / \Delta x)(\partial c / \partial x)]_{j-1}$, and the flux leaving the element, $[A(D_L / \Delta x)(\partial c / \partial x)]_j$, plus advection, $(A \bar{u} c)_j$, across the element, reactions, dc/dt , and rate of mass change due to sources and sinks, S_j . The reactions term includes the changes due to growth (e.g., reaeration) and decay (e.g., deoxygenation). The sources and sinks terms include the transformations (e.g., benthic oxygen demand, volatilization, and bio-

logical transformation) which may add or withdraw the mass internally. The mass of element j may also be externally increased (e.g., from dischargers) or reduced (e.g., from withdrawal), and such change is accounted for by a source or a sink. For steady-state conditions, the accumulation, $\partial M_j / \partial t$, becomes zero. The finite-difference form of Eq. (14) is successively applied to all computational elements of the river system. If any computational element is subjected to an external load, the mass released from that load is added to the system. If the external load is due to discharger m releasing pollutant n , then the mass that is added to the element can be expressed as a linear function of the fractional removal level x_{imn} as

$$q_m c_{mn} = q_m c'_{mn} (1 - x_{imn}) \quad (15)$$

where c_{mn} and c'_{mn} = concentrations of pollutant n after and prior to treatment to a level of x_{imn} , respectively; and q_m = effluent flow from the discharger. Similar to the mass balance, QUAL2E also performs hydraulic and temperature balances in element j .

For each set of fractional removal levels, a simulation run of the water quality model QUAL2E yields a spatial distribution of water quality in terms of indicator concentrations, c_{il} . In water quality management problems, the number of reasonable water quality management alternatives is sufficiently large to preclude simulation of each alternative to select that which is most desirable. Optimization models have been developed and applied as a means of substantially reducing the number of management alternatives that need to be evaluated. In the present work, a GA is used as an optimization tool to solve the fuzzy optimization problem and obtain an optimal solution vector (set of optimal fraction removal levels and minimum satisfaction level) for the WLA problem.

Genetic Algorithms

The fuzzy optimization formulation results in a nonlinear optimization problem when α_{il} and β_{imn} assume values other than unity. The nonlinearity of the constraints and inclusion of an external water quality simulation model for obtaining c_{il} in a river system preclude applicability of linear programming (LP) as an optimization tool. In addition, the disadvantages of conventional nonlinear programming algorithms in discriminating between global and local optimal solutions make them unattractive. GAs, known for assuring a global or near global solution with high probability, are more suitable for the current formulation.

GAs, developed by Holland (1975), are search algorithms based on the mechanisms of natural selection and natural genetics. The first step in applying the GA is representing a candidate solution of the optimization problem with a suitable coding structure, referred to as a chromosome. Elements of a chromosome, known as genes, are juxtapositioned to represent the decision vector. In the WLA problem, the decision vector (X, λ) consists of the set of fraction removal levels, X , and the satisfaction level, λ . Among coding types available for GA, binary and real codings are popular. In the binary coding type, a gene is represented in binary numbers (i.e., 0 or 1), whereas in the real coding type, it is represented by a real number within specified bounds. A set of genes (e.g., 0101) in the binary or a single gene in the real type coding indicates a decision variable (either fraction removal level x_{imn} or minimum satisfaction level λ in the present study) of the optimization. The bounds, the desired level of accuracy, and the constraint handling method influence the design of chromosome coding type. GAs, being an unconstrained optimization technique,

need indirect methods such as decoders or the penalty function method to handle the constraints. In the present work, homomorphous mapping (HM), a decoder-based, indirect method, is chosen to handle the constraints of the fuzzy optimization problem. Details of HM are given in the subsequent section.

HM requires that each gene of the chromosome be represented with a number between -1 and 1 and so the decision vector with $[-1, 1]^{lc}$ cube to convert into a feasible solution. $[-1, 1]^{lc}$ is the Cartesian product of lc sets $[-1, 1]$, where lc is the length (or number of genes) of the chromosome. As required for HM, the real coding type is used for representation of the chromosome in the present study. After processing through HM, the elements of the solution vector represent the set of fraction removal levels, x_{imn} , and the satisfaction level, λ . Suppose a river system consists of three dischargers, each with a single pollutant. Then for a given water quality indicator i , the representation vector includes four variables, viz., three fractional removal levels and a satisfaction level. The upper and lower bounds on the decision variables specified in crisp constraints Eqs. (11) and (12) are implicitly considered in the HM procedure. The fuzzy optimization problem, Eqs. (7)–(12), thus reduces to Eqs. (7)–(10). After receiving the feasible solution vector, the fitness function of the chromosome is evaluated. In a similar way, fitness functions of all the chromosomes of the population are determined. Basic operators including reproduction, crossover, and mutation are applied to obtain a new population of individuals. This cycle of evaluating fitness functions of the population of individuals and applying reproduction, crossover, and mutation to generate offsprings is referred to as a generation. Each generation creates a new population. The GA runs over a fixed number of successive generations. It is hoped that the maximum and average fitness values of the new population will be higher than those of the old population. Monitoring parameters such as maximum and average population fitness values successively over the number of generations provides an indication of how the solution converges toward the optimum. No improvement over a few generations indicates convergence. In a converged population, 95% of all the individuals possess the same fitness function value [De Jong 1975].

The solution corresponding to the highest fitness function in the last generation is the optimal solution with an objective function value equal to the maximum fitness function. Again a fixed number of iterations are performed with a set of GA parameters, such as population size, crossover probability, and mutation probability. In a similar way, different runs are made with different parameter sets to validate the optimal solution. The maximum of all such optimal solutions obtained after performing a number of runs represents the global optimal solution.

HM calls QUAL2E while mapping the coded $[-1, 1]^{lc}$ vector into the decision vector at the time of fitness function evaluation. The next section gives a description of the HM method.

Homomorphous Mapping

The constraint handling methods of GA are grouped (Koziel and Michalewicz 1999) into the following categories: (1) methods based on preserving feasibility of solutions; (2) methods based on penalty functions; (3) methods that make a clear distinction between feasible and infeasible solutions; (4) other hybrid methods that combine evolutionary computation techniques with deterministic procedures for numerical optimization problems; and (5) methods based on decoders. A review of all the constraint handling methods of evolutionary optimization may be found in

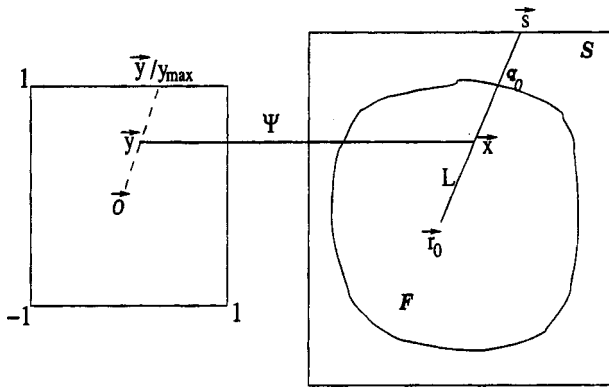


Fig. 3. Mapping Ψ from cube $[-1,1]^{lc}$ into space \mathcal{F} (two-dimensional case)

Michalewicz (1995) and Michalewicz et al. (1996). Among all the constraint handling methods for evolutionary algorithms, methods based on decoders have become popular owing to the following advantages over other methods: (1) they need no additional parameters; (2) there is no need to evaluate or penalize infeasible solutions; (3) a solution located on the edge of the feasible region may be more easily approached; and (4) there is no need for special operators.

In a decoder system a chromosome gives instructions how to build a feasibility space. Each decoder imposes a mapping Ψ between a representation space (e.g., a space of binary strings, vectors of real numbers, and the like) and a feasible part of the solution space \mathcal{F} . It is important that several conditions are satisfied while designing the decoder (Koziel and Michalewicz 1999): (1) for each solution $\vec{x} \in \mathcal{F}$, there is an encoded solution \vec{y} ; (2) each encoded solution \vec{y} corresponds to a feasible solution \vec{x} ; (3) all solutions in \mathcal{F} should be represented by the same number of encodings \vec{y} ; (4) the transformation Ψ is computationally fast; and (5) it has the locality feature in the sense that small changes in the coded solution result in small changes in the solution itself. HM is a decoder-based method, which transforms the lc -dimensional (lc is the size of the decision vector and also the size of the chromosome) cube $[-1,1]^{lc}$ (representation space) into the \mathcal{F} of the optimization problem. Each element of the decision vector is coded with a number between -1 and 1 . For a two element ($lc=2$) decision vector, $[-1,1]^2$ represents a square (Fig. 3).

In the HM, other than the mapping Ψ , an additional one-to-one mapping, $g: [-1,1]^{lc} \rightarrow \mathcal{S}$, between the $[-1,1]^{lc}$ cube and the search space \mathcal{S} of the problem is defined. The mapping g maps the boundary of the $[-1,1]^{lc}$ cube to a boundary point \vec{s} in \mathcal{S} . A reference point, $\vec{r}_0 \in \mathcal{F}$, forms a line segment L with \vec{s} (Fig. 3). If \mathcal{F} is convex, then L intersects the boundary of \mathcal{F} in precisely one point, as shown in Fig. 3. In the case of a nonconvex \mathcal{F} , L intersects the boundary at multiple points (Fig. 4). With the help of the intersection points, HM establishes a one-to-one mapping Ψ between decoded solution \vec{y} ($y_p \in \vec{y}$ and $y_p \in [-1,1]^{lc}$) and feasible point \vec{x} on L in the part of \mathcal{F} . The one-to-one mapping Ψ is defined as

$$g(\vec{y}) = \vec{x} \quad (16)$$

where $x_p = y_p[u(p) - l(p)]/2 + [u(p) + l(p)]/2$, for $p = 1, \dots, lc$; and $u(p)$ and $l(p)$ are upper and lower bounds on decision variable x_p . For $y_p = -1$, the corresponding $x_p = l(p)$ and similarly for $y_p = 1$, $x_p = u(p)$. Each point of \mathcal{S} has a unique point in the

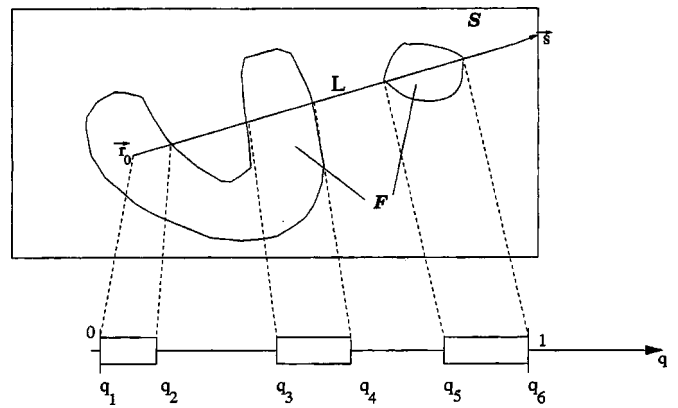


Fig. 4. Line segment in nonconvex space \mathcal{F} and corresponding subintervals (two-dimensional case)

$[-1,1]^{lc}$ cube. If (\vec{y}/y_{\max}) (where $y_{\max} = \max_{p=1, \dots, lc} |y_p|$) is the boundary point of the $[-1,1]^{lc}$ cube, then $g(\vec{y}/y_{\max}) = \vec{s}$ is the relating boundary point of \mathcal{S} (Fig. 3).

A line segment L between any reference point $\vec{r}_0 \in \mathcal{F}$ and \vec{s} is defined as

$$L(\vec{r}_0, \vec{s}_0) = \vec{r}_0 + q(\vec{s} - \vec{r}_0) \quad (17)$$

If \mathcal{F} is convex, L intersects at $q_0 \in [0,1]$, whose point (\vec{x}_{q_0}) is $\vec{r}_0 + q_0(\vec{s} - \vec{r}_0)$. Note that L between $q=0$ and $q=q_0$ falls entirely in \mathcal{F} . For convex \mathcal{F} , HM defines a mapping between $[-1,1]^{lc}$ and the feasible line segment (between 0 and q_0) as

$$\psi(\vec{y}) = \begin{cases} \vec{r}_0 + y_{\max} \cdot q_0 \cdot (\vec{s} - \vec{r}_0) & \text{if } \vec{y} \neq \vec{0} \\ \vec{r}_0 & \text{if } \vec{y} = \vec{0} \end{cases} \quad (18)$$

Fig. 3 shows Ψ for a two-element cube. Because $y_{\max} \in [0,1]$, Ψ results always in a feasible solution \vec{x} . For a nonconvex \mathcal{F} , L may intersect the boundary of \mathcal{F} in many points (Fig. 4). If there are kT subintervals of feasibility, $[q_1, q_2], \dots, [q_{2kT-1}, q_{2kT}]$, the total width of L in the \mathcal{F} region is $d = \sum_{k=1}^{kT} d_k$, where $d_k = q_{2k} - q_{2k-1}$ and $q_{2k-1} < q \leq q_{2k}$. The length of the line segment between \vec{r}_0 and q is

$$\gamma(a) = q_{2k-1} + d_k [a - \delta(q_{2k-1})] / [\delta(q_{2k}) - \delta(q_{2k-1})] \quad (19)$$

where $k = \text{small index such that } a \leq q_{2k-1}$ (i.e., q_{2k-1} is the upper bound of the interval in which a exists); and $\delta(q) : \cup_{k=1}^{kT} (q_{2k-1}, q_{2k}] \rightarrow [0,1]$ is the reverse mapping of $\gamma(a)$, and is given as

$$\delta(q) = \left(q - q_{2k-1} + \sum_{k=1}^{kT-1} d_k \right) / d \quad (20)$$

The HM for a nonconvex \mathcal{F} then can be given by

$$\psi(\vec{y}) = \begin{cases} \vec{r}_0 + q_0 \cdot (\vec{s} - \vec{r}_0) & \text{if } \vec{y} \neq \vec{0} \\ \vec{r}_0 & \text{if } \vec{y} = \vec{0} \end{cases} \quad (21)$$

where $q_0 = \gamma(|y_{\max}|)$. It is identical to the mapping for convex space in Eq. (18). The only difference is in the use of additional mapping γ . The implementation of HM therefore primarily involves the determination of intersection point q_f ($0 \leq f \leq 2kT$).

Simulation-Optimization

The coupling between simulation and optimization allows the advantages of both modules to be retained within a single frame-

work. The method is quite general in its applicability for optimization of a nonlinear, distributed parameter system (Gorelick et al. 1984). Numerical difficulties involved in finite-difference simulation, representing partial differential equations, do not endanger the convergence of the optimization. Such combined models are well suited for water quality decision support systems.

The S-O approach adopted in the present work interfaces QUAL2E, GA, and HM as independent modules to solve the fuzzy optimization problem. The coupling of simulation and optimization comprises two stages: (1) preinterface; and (2) interface. The preinterface stage involves initialization of the individual modules before activating simulation-optimization. Interface involves communication among the modules to solve the fuzzy optimization problem, Eqs. (7)–(12).

Preinterface

A simulation model generally needs a large amount of data for generating the response of the system. This includes data on system configuration, initial and boundary conditions, and other parameters that depend on actual field conditions of the system. Similarly, it is necessary to specify a few parameters and convergence criteria relevant to the optimization component and modify the constraints to be suitable to the particular optimization format. The preinterface stage involves assigning the required data to QUAL2E, GA, and HM and modifying the fuzzy constraints before any interaction among the modules actually begins.

Initialization of QUAL2E

The data required for QUAL2E is grouped into three categories, data regarding the stream system, global variables, and forcing functions. Stream system data consists of details of river discretization; locations of head waters, effluent loads, and junctions; and length of reaches and computational elements. The global variables include simulation settings, such as units (e.g., metric, English), simulation type (e.g., steady state or dynamic), water quality constituents to model (e.g., DO, BOD, temperature, forms of nitrogen, etc.), and temperature correction factors to modify reaeration and deoxygenation rates with respect to actual temperature. Forcing functions are user-specified inputs that drive the system being modeled. These inputs are specified in terms of flow and quality characteristics of both head water and point loads and local climate data. The data falling under these three categories are incorporated into the input file of QUAL2E and remain fixed for all simulations. The input file also consists of the fraction removal levels, which are the decision variables of the fuzzy optimization model. During each call to QUAL2E, the set of fractional removal levels of the dischargers in the input file is replaced with the set provided by HM. Each run of QUAL2E results in the system response in terms of the concentrations of the water quality indicators, or state variables, which are written to an output file. The state variable data required by HM is taken from this QUAL2E output file.

Initialization of Genetic Algorithm and Homomorphous Mapping

The GA needs an assignment of values to parameters such as crossover and mutation probabilities, maximum number of generations, population size, and type of coding, crossover, reproduction, and mutation operators. All individuals (e.g., chromosomes) whose elements (e.g., genes) belong to a $[-1, 1]^{lc}$ cube are gen-

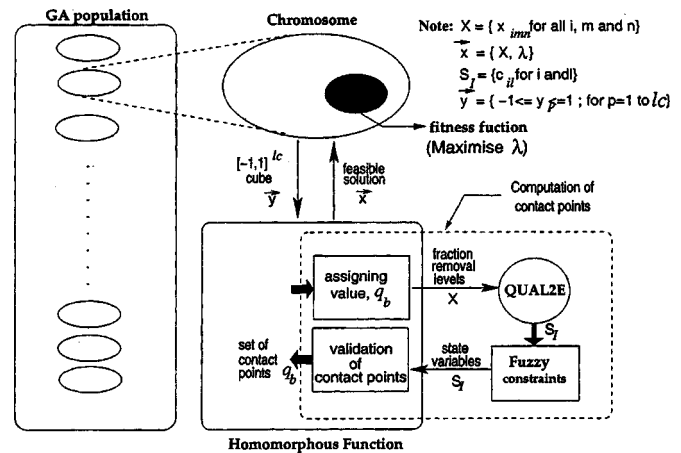


Fig. 5. Integration of QUAL2E, fuzzy optimization model, and genetic algorithm

erated to form the initial population. Each individual is represented as \vec{y} ($-1 \leq y_p \leq 1, \forall p, p=1 \dots lc$). In the present work, PGAPack (David 1996), a library of GA routines, is used for implementing GA as an optimization tool. PGAPack, or the Parallel Genetic Algorithms Package, developed by Argonne National Laboratory, is used to implement the GA operators. PGAPack is a widely used, general purpose, neutral data structure (e.g., suitable for binary, real, integer, and gray codings), parallel GA library. This library contains various operators of selection, crossover, and mutation for various codings. PGAPack ensures implementation of parallel processing in the case of a large number of function evaluations.

HM requires that all constraints be expressed in the form of $g(\vec{x}) < 0$. Accordingly, constraint Eqs. (7)–(12) are altered to fit into the HM module. A feasible initial point, \vec{r}_0 , is required to be specified as a reference point for the HM. HM also needs parameters such as precision and the number of iterations to be performed in finding the intersection points.

The next stage is to develop a communication mechanism between QUAL2E and HM to determine the value of the fitness function.

Interface

The main objective of the interface among GA, HM, and QUAL2E is to evaluate the fitness function of the chromosome abstracted in the form of a $[-1, 1]^{lc}$ cube. Fitness function evaluation of all chromosomes is performed before applying any GA operator (e.g., reproduction, crossover, and mutation) to obtain improved solutions. In general, a feasible solution of the constrained optimization problem is necessary for evaluating the fitness function. Therefore, it is essential to map the chromosome abstracted in the $[-1, 1]^{lc}$ cube, \vec{y} , into a feasible solution, \vec{x} , which represents the fraction removal levels, x_{imn} , and the satisfaction level, λ , to estimate the fitness function. The interface stage indeed focuses on the interaction among the QUAL2E, GA, and HM to convert the abstracted chromosome into a feasible decision vector. This interaction is depicted in Fig. 5. The S-O process starts when the chromosome requests HM, given \vec{y} , to realize the solution vector, \vec{x} . HM makes a number of calls to QUAL2E to get the state vector S_j ($c_{ij} \in S_j$), which is used in feasibility checking. After performing a few interactions, HM returns feasible point \vec{x} .

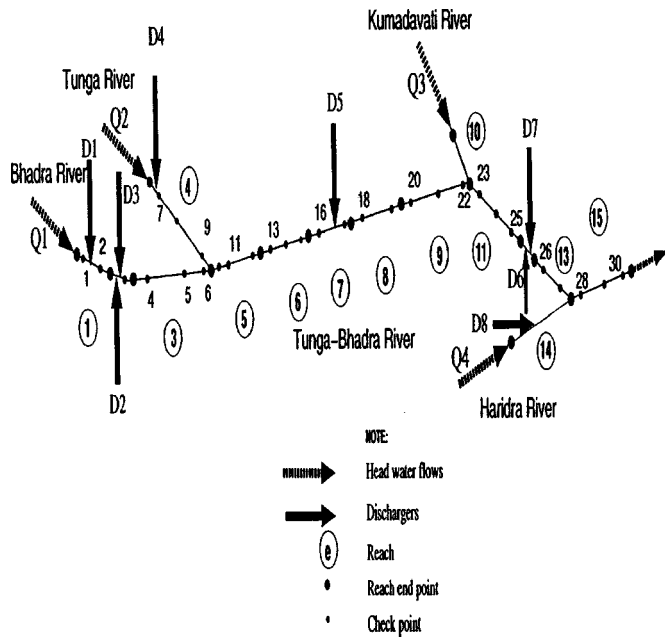


Fig. 6. Schematic of Tunga-Bhadra River system

Eqs. (18) and (21) represent the HM for a convex and nonconvex feasible space \mathcal{F} , respectively. In HM, for a given decoder representation \vec{y} , mapping g defines boundary point \vec{s} of \mathcal{S} . By trail and error or conservative estimate, an initial feasible and reference point \vec{r}_0 can be set without difficulty. The main effort in the implementation of HM is in locating the intersection points q_f of feasible space \mathcal{F} with the line segment L . In the present study, GENOCOP, the genetic algorithm for numerical optimization for constrained problems (Koziel and Michalewicz 1999), is used to handle the constraints in conjunction with PGAPack. Although it also has the capability of implementing GA, in the present case, GENOCOP is used only for mapping.

In GENOCOP, the procedure of finding the intersection points involves expressing each constraint, $g_b(\vec{x})$, as a function β_b of one independent variable, q_b ($0 \leq q_b \leq 1$). The solution of β_b yields the value of q_b , which is the intersection point of the constraint $g_b(\vec{x})$ with the line segment L . In this way, all the intersection points are determined. Among the intersection points, a few points describe the contact points of the boundary of \mathcal{F} and the line segment L . These contact points can be found by checking for the feasibility on both sides of the intersection points on the line segment L . The function β_b is given as

$$\beta_b(q_b) = g_b[L(\vec{r}_0, \vec{s})] = g_b[\vec{r}_0 + q_b(\vec{s} - \vec{r}_0)] \quad \text{for } 0 \leq q_b \leq 1 \quad \text{and } b = 1, \dots, bT \quad (22)$$

In total, there are bT number of constraints g_b excluding the constraints expressing bounds on the decision variables. Each constraint is represented as a function of one independent variable q_b . In order to find the intersection points of g_b with L , the interval $[0,1]$ is partitioned into v subintervals $[v_{e-1}, v_e]$, where $v_e - v_{e-1} = 1/v$ ($1 \leq e \leq v$), so that equations $\beta_b(q_b) = 0$ have at most one solution in every subinterval. The existence of β_b in a subinterval may be found by a binary search. The next step is to check for the feasibility on both sides of the intersection points on line segment L . The presence of feasibility on both sides L indicates that the constraint is within the feasible region and is not the boundary of feasible space \mathcal{F} . Existence of feasibility on one side indicates that the constraint is bounded and the point is also the meeting point between the line segment L and feasible region \mathcal{F} (i.e., $q_b = q_f$). In the process of finding the intersection points, HM calls the simulation model QUAL2E a number of times. The interaction between HM and QUAL2E is as follows:

- In an interval $[v_{e-1}, v_e]$, HM assigns a value to unknown q ($v_{e-1} \leq q \leq v_e$), substitution of which in Eq. (22) yields a solution vector, \vec{x}_{new} , representing the fraction removal levels as well as the satisfaction level.
- The current fraction removal level set, $X_{\text{old}}(\in \vec{x}_{\text{old}})$ of the QUAL2E input file is replaced with the set $X_{\text{new}}(\in \vec{x}_{\text{new}})$. A run of QUAL2E then generates the state variables, c_{il} ($c_{il} \in S_f$), and presents it to the QUAL2E output file.
- The fuzzy constraints of the fuzzy optimization model, Eqs. (7)–(12) are evaluated using the state variable set S_f .
- HM then validates the assigned value to q , based on the feasibility of the fuzzy constraints. Then, a further check is made to ascertain that all the intersection points q are identified.

The intersection points are used to obtain the contact points as discussed earlier. After locating the contact points, q_f , HM applies suitable transformations [Eqs. (18) and (21)] to generate the required decoded feasible decision vector \vec{x} corresponding to \vec{y} and returns it to the chromosome. The fitness function corresponding to the feasible vector \vec{x} is evaluated. The procedure of applying GA operators of reproduction, crossover, and mutation then proceeds as described in the previous section.

Model Application

The S-O methodology is demonstrated through a case study for the control of point BOD waste discharges in the Tunga-Bhadra River in south India. The model is applied to a 180 km reach of the river, which stems from four headwaters—Tunga, Bhadra, Kumadavathi, and Haridra—and eight point loads—five industrial and three municipal effluents. For simplicity, no incremental flow or withdrawal along the stream is assumed to influence the flow in the river system. A schematic of the Tunga-Bhadra River stretch considered for the study is shown in Fig. 6.

Table 1. Headwater Flows Data

River	Gauge station	River flow data			Dissolved oxygen concentration (mg/L)	Biochemical oxygen demand concentration (mg/L)
		Mean (m ³ /s)	Standard deviation (m ³ /s)	Assumed value (m ³ /s)		
Tunga	Shimoga	166.89	351.36	131.75	6.5	1
Bhadra	Bhadra Dam	17.80	78.96	17.80	6.5	1
Kumadvathi	Kuppalur	14.94	39.17	11.02	6.5	1
Haridra	Byaladahalli	13.90	34.32	10.47	6.0	1

Table 2. Effluent Flow Data

Discharger number	Location		Biochemical oxygen demand concentration (mg/L)	Dissolved oxygen concentration (mg/L)	Effluent flow data		
	Reach number	Element number			Mean (m ³ /s)	Standard deviation (m ³ /s)	Assumed value (m ³ /s)
1	1	2	1,000	2.0	0.705	0.154	1.167
2	2	1	440	2.0	0.308	0.077	0.539
3	2	2	300	2.0	0.026	0.002	0.032
4	4	2	900	2.0	0.436	0.109	0.763
5	7	9	222	2.0	0.024	0.006	0.042
6	12	1	600	2.0	0.129	0.032	0.225
7	12	2	450	2.0	0.952	0.240	1.672
8	14	2	900	2.0	0.867	0.216	1.515

In QUAL2E, a reach is defined as a stretch of river along which model input parameters or coefficients (physical, chemical, and biological) remain constant. A new reach is defined at locations where a new junction is encountered or a significant change in the model input coefficients occurs. Accordingly, the 180-km-long stretch of the river is divided into 15 reaches of varying lengths, each of which is further discretized into computational elements of 1 km in length. The maximum length of a reach is restricted to 20 km, following the QUAL2E restriction of twenty computational elements within each reach.

All eight dischargers discharge BOD effluents to the river, and the water quality indicator of interest is DO at 30 checkpoints due to these point sources of BOD, as shown in Fig. 6. A maximum number of three checkpoints are considered in a reach. The data required for the simulation model have been collected from government agencies (Water Resources Development Organization, Bangalore, Karnataka Pollution Control Board) and a few industries discharging their effluents in the stream. The data are processed and analyzed in an earlier work (Suresh 1999). This data includes: (1) details of discharges, including DO and BOD concentrations of both head water and effluent flows; (2) details of river geometry; and (3) temperature data for the waste load and river flow. Statistical characteristics of the river and effluent flows, along with the DO and BOD concentrations, are given in Tables 1 and 2. A trapezoidal cross-sectional shape with side slopes of 1:1 is considered for the river. The minimum permissible level with respect to DO is set to 4 mg/L for a portion of the Bhadra River (reaches 1–3) and 6 mg/L for the remainder of the river. The desirable level of DO is set to the saturation level of DO (i.e., $c_{il}^D = DO_{sat}$, $\forall i, l$), which is computed in QUAL2E based on temperature data. A minimum fractional removal level of 35% and a maximum level of 90% are assumed for the dischargers.

Because the sets of pollutants and water quality indicators contain only one element each (i.e., BOD and DO, respectively), the suffixes n and i are dropped from the constraints, for convenience. The river system contains 30 water quality checkpoints ($l=30$) and eight dischargers ($m=8$). Denoting the DO at the water quality checkpoint l by c_l , and the fraction removal level for the discharger m by x_m , and using linear membership functions for

the fuzzy goals (i.e., $\alpha_{il} = \beta_{imn} = 1$), the fuzzy waste load allocation model, originally expressed in Eqs. (7)–(12), is written as

$$\text{Maximize } \lambda \quad (23)$$

subject to

$$[(c_l - c_l^L)/(c_l^D - c_l^L)] \geq \lambda \quad \forall l \quad (24)$$

$$[(x_m^M - x_m)/(x_m^M - x_m^L)] \geq \lambda \quad \forall m \quad (25)$$

The fuzzy constraint in Eq. (8) is satisfied only when the concentration, c_l , is within its limits as specified in crisp constraint of Eq. (10). Design of the chromosome in the application of the GA ensures that the elements of the decision vector are within their bounds as specified in crisp constraints of Eqs. (11) and (12). Therefore, the crisp constraint Eqs. (10), (11), and (12) become redundant and are omitted from the optimization formulation.

The DO concentration at checkpoint l , c_l , is obtained from QUAL2E for a given set of fraction removal levels, x_m , and other fixed data. It may be noted that the DO at a checkpoint is affected only by the effluents, controlled by the fraction removal levels, upstream of that checkpoint.

Results and Discussion

Implementation

The S-O methodology, which integrates the GA, HM, and QUAL2E, is implemented on a SUN ULTRA-60 workstation, Sparc series, utilizing a UNIX OS. As previously explained, the GA, HM, and QUAL2E act as individual modules with interactive communication among them. QUAL2E, being primarily a DOS-based public domain software, requires a few modifications for porting it on a UNIX-based computer system. Additionally, for reasons of computation speed and flexibility to work on different computer systems in the network, the QUAL2E input and output operations have been modified, without affecting the core process of the model.

A typical coding of the decision vector of the optimization problem is shown in Table 3. The table also contains the corre-

Table 3. Coded and Decoded Solution Vector in a Communication between Genetical Algorithm and Homomorphous Mapping

Decision variable	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	λ
Coding of solution vector in form of $[-1, 1]^{lc}$ cube	0.7	0.8	-0.3	0.27	0.75	0.11	-0.67	-0.96	0.1
Corresponding feasible solution (%)	85.7	85.8	78.8	35.0	35.0	45.24	81.5	46.9	0.1

Table 4. Variation of Decision Vector over Generations

Generation number	Fraction removal levels (%)								Satisfaction level, λ
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	
0	74.58	74.57	66.48	35.00	35.00	35.00	35.00	35.00	0.0001
1	75.10	71.92	66.74	38.97	62.37	49.04	38.08	41.63	0.2590
10	74.71	73.83	63.36	46.14	35.00	46.42	39.64	41.29	0.2750
20	74.68	74.69	70.71	52.78	37.29	52.74	35.21	39.49	0.2780
30	74.68	74.39	74.08	52.84	48.96	52.81	39.74	41.75	0.2780
40	74.68	74.39	60.32	52.84	43.48	46.58	35.18	47.07	0.2780
50	74.67	74.40	66.67	52.55	40.15	36.99	39.67	46.06	0.2780
60	74.67	74.40	61.30	52.55	40.15	41.70	39.58	46.06	0.2780
70	74.67	74.40	67.15	52.55	40.15	38.93	39.19	48.77	0.2780
80	74.59	74.40	72.97	52.63	38.36	38.63	45.62	48.51	0.2790
90	74.60	74.39	71.13	53.22	44.13	42.11	45.97	39.80	0.2790
100	74.56	74.39	72.34	53.24	44.07	50.72	41.47	39.81	0.2800

sponding decoded solution that is sent by HM to the GA. A few interactions are performed between HM and QUAL2E in converting the $[-1,1]^{lc}$ cube to a feasible solution. In a typical interaction between the two modules, HM writes the set of fraction removal levels to the input file of QUAL2E. The simulation run of QUAL2E generates and writes the data to an output file, which provides the DO concentrations for the checkpoints of the river system in order to evaluate the constraints pertaining to water quality indicators. Following the interactions between HM and QUAL2E, HM directs the feasible solution to the GA, where routine GA operators are applied.

The selection of parameter values and operators of the GA play a prominent role in enhancing the performance of the GA. Multiple runs with various values of parameters and operator types are made for determining parametric values (e.g., probability of crossover, p_c ; probability of mutation, p_m ; population size; number of generations) and types of operators (e.g., methods for selection, crossover, and mutation) for achieving the global optimum solution. After carrying out multiple runs of the GA with different parameter values, it is found that a p_c of 0.75 with a uniform type crossover, a p_m of 0.122 with mutation range type, tournament selection, and a population size of 200 are appropriate in producing the global solution. Elitism has been used to retain 40% (i.e., 80) of the best individuals of the previous generation in the present generation to prevent loss of the best individuals of a population.

Results

Table 4 and Fig. 7 show progress of the solution across generations. Note that an initial satisfaction level of 0.0001 is improved to a value of 0.278 in only 35 generations. A similar rapid improvement is also observed in maximum and average fitness values. This trend indicates the potentiality of the GA to converge rapidly in a few generations. Although an improvement is witnessed in both the best and the average satisfaction levels, the optimal policy (i.e., fraction removal levels) obtained in generation 100 is not significantly different with respect to the initial solution. Two reasons can be identified as the cause. First, the GA is primarily a fitness function driven optimization technique. The candidacy of a solution is judged as best merely based on its fitness function value, not on the solution composition. Second, the FWLAM formulation aims at maximizing the minimum satisfaction level of the system, and this could imply the existence of multiple solutions. Selection of the best parameters and the best

operator types to use within the GA merely ensures a high resulting fitness value and may not guarantee a favorable fraction removal level policy unless the fraction removal levels are also included in the fitness function evaluation. In the case study, most of the multiple runs of the GA yielded a satisfaction level near 0.2801, but with different optimal policies (e.g., treatment levels to dischargers). This indicates the existence of multiple solutions corresponding to the optimal objective function value.

Selection of the initial point of HM could affect the final solution. To achieve different sets of solutions, the best solution obtained from a run of the GA is provided as the initial point for a subsequent run of the GA. In this way, multiple runs are carried out to obtain a favorable solution. From all of the solutions obtained with respect to the maximum satisfaction level, $\lambda^*=0.2802$, the solution that assigns minimum treatment levels to dischargers is taken as the optimal solution, which is given in Table 5. Highest treatment levels of 74.6% are assigned to dischargers D_1 and D_2 in the system. A high treatment level results for discharger D_1 due to its higher BOD mass (i.e., product of effluent flow and BOD concentration) as compared with other dischargers in the system. Though dischargers D_2 and D_3 release much less BOD mass as compared with D_1 , these dischargers are also assigned high treatment levels because of their proximity to

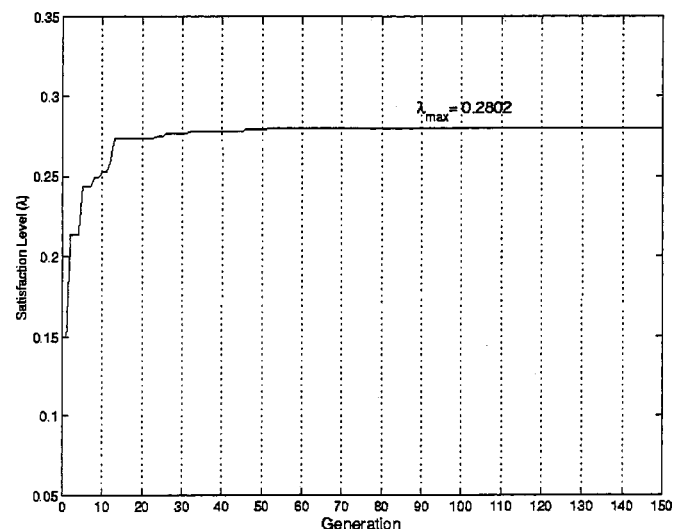
**Fig. 7.** Performance of fitness functions over generations

Table 5. Optimal Fraction Removal Levels

Discharger	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8
Fraction removal (%)	74.6	74.6	66.5	35.0	35.0	35.00	35.0	45.0

discharger D_1 . The high flow from the Tunga head water helps to increase the assimilative capacity and thus improves the DO concentration in the river, resulting in a minimum treatment level to discharger D_4 . The next highest treatment level after D_1 , D_2 , and D_3 is assigned to D_8 . Again, the reason for a high treatment level may be attributed to the high BOD mass D_8 releases into the river system. Other dischargers D_5 , D_6 , and D_7 are assigned the minimum treatment level, 35%.

The upper and lower bounds of λ reflect two extreme scenarios in the system. The upper bound, $\lambda=1$, indicates that all the goals have been completely satisfied and, therefore, represents a no-conflict scenario. The lower bound, $\lambda=0$, indicates that at least one goal has a zero satisfaction level and therefore represents a total conflict scenario. Any intermediate value of λ represents the degree of conflict that exists in the system. The present formulation is aimed at achieving a fair-compromise solution by addressing the degree of conflict in the system. It may be expected for a water quality management problem that the value of the minimum satisfaction level, λ , will be closer to zero than to unity, as demonstrated in the case study. This indicates that a conflict scenario generally cannot be avoided. The existence of a conflict scenario in a water quality management problem is due to the compound effect of the conflicting objectives of the PCA and dischargers and the relatively low assimilative capacity of the river network with respect to the pollutant loads.

Concluding Remarks

This paper presents a simulation-optimization procedure for obtaining fraction removal levels for point sources of pollutants in a river system. A fuzzy optimization model is integrated with a water quality simulation model and genetic algorithms are used for solving the optimization model. To handle constraints in the optimization model, the HM, a decoder-based method for GAs, is used. The paper demonstrates the computational aspects of the integration with a simplified river system in which only the BOD point sources are considered and the DO at a finite number of checkpoints is used as the water quality indicator.

The fuzzy optimization model addresses uncertainty due to imprecision and subjectivity in the water quality management problems. Goals of the pollution control agencies and the dischargers are expressed as fuzzy sets in the model with appropriate membership functions. In a majority of waste load allocation models, the measure of performance is related to the overall cost of pollution control, including the waste treatment cost. Incorporating the cost functions directly into the optimization models poses difficulties due to uncertainty, lack of adequate data, and nonlinearity associated with the cost functions. The fuzzy waste load allocation model has the advantage that the cost functions are eliminated while the goals of the dischargers are incorporated through fuzzy membership functions. The membership functions themselves are subjective statements of the perceptions of the decision makers. The lower and upper bounds, viz., c_{il}^L , c_{il}^D , x_{imn}^L , and x_{imn}^M , used in the membership functions as well as the shape of the membership functions depend on the responses of the particular decision maker to water quality and fraction re-

moval levels. To address uncertainty in these bounds, the fuzzy membership functions themselves may be treated as fuzzy in the model and may be modeled using gray numbers. However, as the emphasis in this paper has been on integrating the HM with GAs and a fuzzy water quality control model, only one level of fuzziness has been considered to reduce complexity in the model. Apart from uncertainty due to fuzziness, a water quality management problem is also characterized by uncertainty due to randomness in stream flow, waste loads, temperature, non-point sources of pollution, water quality parameters, and other variables, along with model uncertainties. Inaccuracies and uncertainties in water quality model outputs also arise most often from the quantity and quality of the data used. Addressing these forms of uncertainties simultaneously with fuzziness should provide a useful direction for future research.

The model formulation and the procedure developed for integration of GA, water quality simulation, and HM are general in nature, and any well-tested water quality simulation model may be used in the procedure. In Fig. 5, in the block showing QUAL2E, any other water quality simulation model may be substituted without altering the framework of the procedure. Numerical stability of the simulation model used in the S-O procedure is an important criterion for choice of the simulation model. If repeated instability keeps occurring in the simulation model, it may not be possible to compute the state variables, and then the convergence of the optimization algorithm may not be ensured. It is therefore necessary that a reasonably stable simulation model be used in the S-O procedure. In the case studied in this paper, difficulties due to numerical instability in the simulation model were not encountered. For the simplified case presented in the paper, because it is a noncoupled, steady-state system with only BOD and DO as the water quality indicators of interest, for example, the S-P model may be used equally effectively in place of QUAL2E, and with linear fuzzy membership functions for all goals, the fuzzy optimization model may be formulated as a linear programming problem (e.g., Sasikumar and Mujumdar 1998; Mujumdar and Sasikumar 2002). Use of QUAL2E and other, more detailed, water quality simulation models along with GA and HM will be advantageous with nonlinear fuzzy membership functions or when coupled systems with interactions between algae, phosphorus, nitrogen, and DO are considered. The procedure presented in the paper provides a framework for such general problems.

Acknowledgment

The writers sincerely thank the two anonymous reviewers for reviewing the manuscript and providing critical comments to improve the paper.

Notation

The following symbols are used in this paper:

- A_j = cross-sectional area of computational element j ;
- a = constant used in HM;
- bT = total number of constraints;
- c_{il} = concentration of water quality indicator i at checkpoint l ;
- c_{il}^D, c_{il}^L = desirable and permissible levels for c_{il} ;
- c_j = concentration of constituent in computational element j ;

c_l = concentration of DO at checkpoint l ;
 c_l^D, c_l^L = desirable and permissible levels for c_l ;
 c_{mn}, c'_{mn} = concentration of pollutant n before and after treatment from discharger m ;
 D_{L_j} = longitudinal dispersion coefficient in computational element j ;
 E_l, E_{il} = fuzzy goal of pollution control agency with respect to c_l and c_{il} , respectively;
 \mathcal{F} = feasible space;
 F_{imn} = fuzzy goal of discharger m with respect to pollutant n to control water quality indicator i ;
 $f_{ilmn}(\cdot, \cdot)$ = transfer function;
 g = mapping between $[-1, 1]^{lc}$ cube and search space \mathcal{S} ;
 g_b = inequality constraint b ;
 kT = number of feasible subintervals on line segment L ;
 L = line segment joining \vec{r} and \vec{s} ;
 L_{ilmn} = concentration of pollutant n prior to treatment from discharger that affects water quality indicator i at checkpoint l ;
 lc = length of chromosome;
 $l(p)$ = lower bound on element p in chromosome;
 M_j = mass of constituents in computational element j ;
 M_u = total number of dischargers located upstream of checkpoint l ;
 N_u = total number of pollutants released from M_u dischargers upstream of checkpoint l ;
 p_c = probability of crossover;
 p_m = probability of mutation;
 Q_j = rate of inflow to element j ;
 q_b = intersection point between line segment L and constraint g_b ;
 q_f = intersection point between line segment L and feasible space \mathcal{F} ;
 q_m = discharge from external source m ;
 q_0 = distance of feasible boundary from \vec{r}_0 on L ;
 q_{2k-1}, q_{2k} = end points of feasible subinterval k on L in HM;
 \vec{r}_0 = reference (also feasible) point belongs to \mathcal{F} in HM;
 \mathcal{S} = search space;
 \mathbf{S}_l = set of concentration levels c_{il} ;
 S_j = external sources or sinks changing mass in computational elements j ;
 \vec{s} = boundary point of search space \mathcal{S} in HM;
 t = time;
 u_j = mean velocity of flow in computational element j ;
 $u(p)$ = upper bound on element p in chromosome;
 V_j = volume of computational element j ;
 v = number of subintervals $[v_{e-1}, v_e]$ in HM;
 v_e = end points of subintervals in HM;
 X = set of fraction removal levels;
 X_{old}, X_{new} = old and new sets of fraction removal levels in typical interaction between HM and QUAL2E;
 \vec{x} = feasible decision vector of fractional removal levels and satisfaction level corresponding to chromosome \vec{y} ;

x_{imn} = fraction removal of pollutant n from discharger m to control water quality indicator i ;
 x_{imn}^L, x_{imn}^M = aspiration and maximum acceptable fraction removal levels set by discharger m ;
 x_{imn}^{MAX} = technologically possible maximum fraction removal level;
 x_{imn}^{MIN} = minimum fraction removal level imposed by pollution control agency;
 x_m = fraction removal of pollutant for discharger m ;
 x_m^L, x_m^M = minimum and maximum fraction removal levels of pollutant set by discharger m ;
 $\vec{x}_{old}, \vec{x}_{new}$ = old and new decision vectors in typical interaction between HM and QUAL2E;
 x_p = member of feasible decision vector \vec{x} ;
 \vec{y} = chromosome whose elements (genes) belong to set $[-1, 1]$;
 y_{max} = absolute maximum element of \vec{y} ;
 y_p = member of \vec{y} ;
 α_{il} = exponent in membership function of fuzzy goal E_{il} ;
 β_b = constraint b expressed as function of independent variable q_b in HM;
 β_{imn} = exponent in membership function of fuzzy goal F_{imn} ;
 γ = mapping between $[0, 1]$ and union of feasible subintervals;
 Δx_j = length of computational element j ;
 δ = mapping between union of feasible subintervals and $[0, 1]$;
 λ = minimum satisfaction level;
 $\mu_{E_{il}}(\cdot)$ = fuzzy membership functions of PCA's goal;
 $\mu_{F_{imn}}(\cdot)$ = fuzzy membership functions of discharger's goal; and
 ψ = homomorphous mapping.

Subscripts

b = constraint index;
 e = subinterval end in HM implementation;
 f = intersection point index in HM implementation;
 i = water quality indicator index;
 j = computational element of reach;
 k = feasible interval index in nonconvex \mathcal{F} ;
 l = checkpoint index;
 m = discharger index;
 n = pollutant index;
 p = gene or element of chromosome; and
 u = upstream of checkpoint l .

References

- Brown, L. C., and Barnwell, T. O., Jr. (1987). "The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and user manual." *Rep. No. EPA/600/3-87/007*, U.S. Environmental Protection Agency, Athens, Ga.
- Burn, D. H. (1989). "Water-quality management through combined simulation-optimization approach." *J. Environ. Eng.*, 115(5), 1011–1024.
- Burn, D. H., and Lence, B. J. (1992). "Comparison of optimization formulations for waste-load allocations." *J. Environ. Eng.*, 118(4), 597–612.
- Burn, D. H., and Yulianti, J. S. (2001). "Waste-load allocation using genetic algorithms." *J. Water Resour. Plan. Manage.*, 127(2), 121–129.
- Camp, T. R. (1963). *Water and its impurities*, Reinhold, New York.

- Canale, R. P. (1976). *Modeling biochemical processes in aquatic ecosystems*, Ann Arbor Science, Ann Arbor, Mich.
- Cardwell, H., and Ellis, H. (1993). "Stochastic dynamic programming models for water quality management." *Water Resour. Res.*, 29(4), 803–813.
- Carmichael, J. J., and Strzepek, K. M. (2000). "A multiple-organic-pollutant simulation/optimization model of industrial and municipal wastewater loading to a riverine environment." *Water Resour. Res.*, 36(5), 1325–1332.
- Dai, T., and Labadie, J. W. (2001). "River basin network model for integrated water quantity/quality management." *J. Water Resour. Plan. Manage.*, 127(5), 295–305.
- David, L. (1996). "Users guide to the PGAPack parallel genetic algorithms library." *Argonne National Laboratory ANL-95/18*, (http://www-fp.mcs.anl.gov/CCST/research/reports_pre1998/comp_bio/stalk/pgapack.html).
- De Azevedo, L. G., Gates, T. K., Fontane, D. G., Labadie, J. W., and Porto, R. L. (2000). "Integration of water quantity and quality in strategic river basin planning." *J. Water Resour. Plan. Manage.*, 126(2), 85–97.
- De Jong, K. A. (1975). "An analysis of the behavior of a class of genetic adaptive systems." PhD thesis, Univ. of Michigan, Ann Arbor, Mich.
- Dobbins, W. E. (1964). "BOD and oxygen relationships in streams." *J. Sanit. Eng. Div., Am. Soc. Civ. Eng.*, 90(3), 53–78.
- Gill, P. E., Murray, W., and Wright, M. H. (1981). *Practical optimization*, Academic, New York.
- Gorelick, S. M., Voss, C. I., Gill, P. E., Murray, W., Saunders, M. A., and Wright, M. H. (1984). "Aquifer reclamation design: The use of contaminant transport simulation combined with nonlinear programming." *Water Resour. Res.*, 20(4), 415–427.
- Han, K. Y., Kim, S. H., and Bae, D. H. (2001). "Stochastic water quality analysis using reliability method." *J. Am. Water Resour. Assoc.*, 37(3), 695–708.
- Harleman, D. R. F. (1978). "A comparison of water quality models of the aerobic nitrogen cycle." *Research Memorandum RM-78-34*, International Institute for Applied Systems Analysis, Laxenburg, Austria.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, Mich.
- Koziel, S., and Michalewicz, Z. (1999). "Evolutionary algorithms, homomorphous mappings, and constrained parameter optimization." *Evol. Comput.*, 7(1), 19–44.
- Maier, H. R., Lence, B. J., Tolson, A., and Foschi, O. (2001). "First-order reliability method for estimating reliability, vulnerability, and resilience." *Water Resour. Res.*, 37(3), 779–790.
- McKinney, D. C., and Lin, K. D. (1994). "Genetic algorithm solution of groundwater management models." *Water Resour. Res.*, 33(6), 1897–1906.
- Michalewicz, Z. (1995). "Genetic algorithms, numerical optimization, and constraints." *Proc., 6th Int. Conf. on Genetic Algorithms*, Morgan Kaufmann, San Mateo, Calif., 151–158.
- Michalewicz, Z., Nazhiyath, G., and Michalewicz, M. (1996). "A note on usefulness of geometrical crossover for numerical optimization problems." *Proc., 5th Annual Conf. on Evolutionary Programming*, MIT Press, Cambridge, Mass., 305–312.
- Mujumdar, P. P., and Sasikumar, K. (2002). "A fuzzy risk approach for seasonal water quality management of a river system." *Water Resour. Res.*, 38(1), 5-1–5-9.
- Murtagh, B. A., and Saunders, M. A. (1980). "MINOS/AUGMENTED user's manual." *Systems Optimization Laboratory Technical Rep. 80-14*, Dept. of Operations Research, Stanford Univ., Stanford, Calif.
- Murtagh, B. A., and Saunders, M. A. (1982). "A projected Lagrangian algorithm and its implementation for sparse nonlinear constraints." *Mathematical Programming Study 16*, North-Holland, Amsterdam.
- Oliveira, R., and Loucks, D. P. (1997). "Operating rules for multireservoir systems." *Water Resour. Res.*, 33(4), 839–852.
- Robert, B. A., Tim, A. W., Connolly, J. P., and Robert, W. S. (1988). "WASP4: A hydrodynamic and water quality model—Model theory, user's manual, and programmer's guide." *Rep. EPA/600/3-87/039*, Environmental Research Laboratory, U.S. Environmental Protection Agency, Athens, Ga.
- Sakarya, A. B. A., and Mays, L. W. (2000). "Optimal operation of water distribution pumps with water quality considerations." *J. Water Resour. Plan. Manage.*, 126(4), 210–220.
- Sasikumar, K., and Mujumdar, P. P. (1998). "Fuzzy optimization model for water quality management of river system." *J. Water Resour. Plan. Manage.*, 124(2), 79–84.
- Sasikumar, K., and Mujumdar, P. P. (2000). "Application of fuzzy probability in water quality management of a river system." *Int. J. Syst. Sci.*, 31(5), 575–591.
- Streeter, H. W., and Phelps, E. B. (1925). "A study of the pollution and natural purification of the Ohio River. III: Factors concerning the phenomena of oxidation and reaeration." *Public Health Bulletin No. 146*, U.S. Public Health Service, Washington, D.C.
- Suresh, H. R. (1999). "A fuzzy optimization model for waste load allocation in streams." MS thesis, Dept. of Civil Engineering, Indian Institute of Science, Bangalore, India.
- Suresh, H. R., and Mujumdar, P. P. (1999). "A neural network model for waste load allocation in rivers." *Proc., Civil and Environmental Engineering Conf.: New Frontiers and Challenges*, Asian Institute of Technology, Bangkok, Thailand, 1(2), 97–104.
- Takyi, A. K., and Lence, B. J. (1994). "Incorporating input information uncertainty in a water quality management model using combined simulation and optimization." presented at *Proc., Int. UNESCO Symp. on Water Resources Planning in a Changing World*, Karlsruhe, Germany.
- Takyi, A. K., and Lence, B. J. (1999). "Surface water quality management using a multiple-realization chance constraint method." *Water Resour. Res.*, 35(5), 1657–1670.