TR(BR) 5/96-97

ASSESSMENT OF PROBABILITY DISTRIBUTION OF DISSOLVED OXYGEN

NATIONAL INSTITUTE OF HYDROLOGY
JAL VIGYAN BHAWAN
ROORKEE - 247 667
1996-97

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

It has been recognized that processes in natuarl stream environ ments are inherently random. The ability to quanitify the probabilistic status of stream environments is of vital importance in water quality management decision making. This report presents a methodology to assess the probability distribution of dissolved oxygen based Streeter-Phelp equation. The methodology involves the use of First-order analysisand Monte Carlo simulation to analyze the uncertainty associated with dissolved oxygen. The mean and standard deviation of the parameters hvae been assumed same for studied probability distributions.

The critical DO is found to be approximately normally distributed. The mean value of DO at the critical location is found to be more or less same, irrespective of the method used. But the level of uncertainty associated with the DO is found to be considerably different. The Monte Carlo simulations with log-normally dsitributed input is found to give the lower uncertainty in the DO levels.

For estimation of travel time, Monte Carlo simulation with lognormally distributed input variables is found to be a preferable method. For the over all DO profile in the stream, First order analysis predicting the same DO level as given by Monte Carlo simulationand lower level of uncertainty, is found to be more justified

1.0 INTRODUCTION

A major portion of the complexity associated with water quality modelling and prediction is the inherent randomness exhibited through the stream environment. Not only are the physical and biological processes not clearly defined, but an imposing number of uncertainties are associated with the various processes occurring with in the stream environment. Several researchers have already attempted to analyze these uncertainties. Loucks and Lynn (1966) investigated the effect of inherent uncertainty due to the natural variations in stream flow and waste flow on the probability distribution of dissolved oxygen (DO), Padgett and Rao (1979) presented a joint probability distribution for biochemical oxygen demand (BOD) and (DO), and Kothandaraman and Ewing (1969) and Chadderton et al. (1982) have investigated the effect of stochastic nature of the model parameters in assessing the probability distribution of DO deficit. In achieving effective environmental control, the procedure of Waste Load Allocation (WLA) should consider the natural inherent randomness of water quality parameters. The allocation process involves the estimation of stream assimilative capacity. Characterisation of point and diffuse source inputs, reserve capacity allocation, and a subsequent assignment of available capacity to designated discharges. Procedural steps require a determination of the Total Maximum daily load (TMDL) and a distribution of assimilative capacity in an equitable manner.

Realising the existence of such uncertainties in water quality modelling, the prediction of DO deficit or DO concentration with in a given reach of stream is no longer a deterministic exercise. Rather, the DO deficit must be treated as a random variable. In probabilistic water quality analysis, it is typical to deal with the problems of assessing the probability of water quality violation. To perform such probability computations,

knowledge about the statistical properties and the distribution of water index must be known.

Although significant research has been conducted in the uncertainty analysis of stream DO, most of these studies have been concerned with the variability of DO concentrations due to model parameter uncertainty (Kothandaraman and Ewing 1969; Hornberger 1980; Chadderton et al. 1982). However, there have been some attempts to derive analytical expressions for the exact probability distribution associated with the DO deficit. Thayer and Krutchkoff (1967) utilized aq stochastic birth and death process to obtain an expression for the probability distribution of DO deficit. Thayer and Kurtchkoff (1967) utilized a stochastic birth and death process to obtain an expression for the probability distribution of DO concentration without considering the uncertainties of the model parameters. Esen and Rathbun (1967) assumed the reaeration and deoxygenation rates to be normally distributed and investigated the probability distribution for DO and BOD using a random walk approach. Notably, Padgett et al. (1977) developed a joint probability density function for the BOD and DO concentrations for solving a stochastic differential equation, and Padgett and Rao (1979) later developed a non-parametric probability density function of BOD and DO.

From a practical viewpoint, the main disadvantage of each of the aforementioned methods is that the resulting probability distributions derived for the DO deficit are very complicated. The required mathematical skills needed for such sophisticated approaches would make it difficult for most engineers to apply. Furthermore, all these analytically derived probability distribution functions for the DO can only be obtained by using very simple distributions for the model parameters such as uniform and normal. When

distributions other than those simple ones are used to describe the randomness of water quality parameters (which could well be the case in reality), the analytical derivation of a probability distribution for the DO would be extremely difficult, if not impossible.

Another approach that is frequently applied by engineers is the Monte Carlo simulation. The method has recently been incorporated into the enhanced QUAL2 model, called QUAL2-UNCAS, by Brown and Barnwell (1987). This brute force enumeration scheme requires a large number of repetitions, which could be computationally expensive. Ofcourse, with the advent of computing power and efficiency of computers, the weight of such concern will be gradually diminishing. However, at the present time, computational efficiency and cost remain an important concern in practical engineering problem solving.

In support of a more tractable methodology, this study examines an approximate approach to probabilistic water quality analysis in that the statistical moments of the DO deficit are estimated by the first order analysis. The statistical moments estimated are then incorporated with an appropriate probability distribution model for the DO deficit. However, practical issues that can be raised are: (1) Which is an appropriate probability model for the DO deficit? and (2) if there is one such probability model , how sensitive is it to the distribution of water quality parameters? Theoretically, the exact probability distribution of the DO deficit should change if the distribution of water quality parameters is changed. However, from a practical view point, it is worthwhile to investigate the appropriateness of some commonly used probability models in describing the random characteristics of the DO deficit computed by the Strata-Phelps equation. Because the distribution of the DO deficit may in theory be dependent on the statistical properties of the water quality parameters themselves. The candidate probability distributions included in

the study were normal, and lognormal. To characterize these distribution completely, various statistical parameters must be known. To do this, the mean and variance of the DO deficit were estimated using first order uncertainty analysis. This information was then used to compute the appropriate statistical parameters for each of the candidate probability models using the moment-parameter-relationships that can be found in Hasting and Peacock (1974) and Patel et al. (1976).

2.0 UNCERTAINTY ANALYSIS

Over the years, a progression in the modelling of water quality phenomena has developed to simulate chemical, physical and biological processes occuring in river waters. Their possible applications range from identifying instreaming processes affecting river water quality to forecasting the quality for operational purposes. At the outset, deterministic models were utilized. However, a fundamental characteristic of environmental engineering phenomena is their intrinsic stochastic nature. In addition to the frequent problem of having an inadequate number of observations, the problem of modelling the uncertainty in the observations or environmental variables resulting from measurement errors and other disturbances of natural origin is a challenging problem.

Certainly, deterministic models are much simpler to deal with mathematically in contrast to the stochastic models. However, deterministic models can predict only the mean or expected valuue of the process or worst-case situation and this is an acceptable solution for modelling processes having very small uncertinty. The uncertainty associated with water quality processes (attributed due to a number of reasons) is not only complicated but also very high in degree.

There is always some uncertinty, both in the evaluation of field data and in the use of mathematical models to predict the outcome of natural processes are still not completely understood and the full representation is usually complicated and too costly to implement. There is also some inherent variability and randomness in natural processes and their measurements. The initial conditions may also be random, either because measurements are biased by random variations. The model coefficients are random eithe because our assessment is not prefect or because of random variations in measurements. Inputs may

also be uncertain because estimates of future loadings, based on projections and future wastewater technologies, may be biased.

With these types of considerations in mind, the intent in the following is to examine some of available stochastic modelling approaches. In the present study, the uncertainty is considered implicitly with the dissolved oxygen model using first-order analysis and Monte Carlo simulations.

2.1 FIRST ORDER ANALYSIS OF UNCERTAINTY

The method of first order uncertainty analysis can be used to estimate the amount of uncertainty, or scatter, of a dependent variable due to uncertainty about the independent variables included in a functional relationship. The method is applied with the assumption that all covariances among variables are zero. First order uncertainty analysis has been described in detail by Benjamin and Cornell (1970). Example applications have also been presented by Burgess (1979) and Chadderton and Miller (1980).

The use of first - order uncertainty analysis is popular in all fields of engineering because of its relative ease in application to a wide array of problems. The detailed theory and mathematics of first order uncertainty analysis can be found in Benjamin and Cornell (1970) and Cornell (1972). As an example of such use in the water quality field, Burges and Lettenmaier (1975) have utilised the method to investigate the uncertainty in predictions of BOD and DO with in a stochastic stream environment.

Essentially, first order analysis provides a methodology for obtaining an estimate for the moments of a random variable which is a function of one or several random variables. It estimates the uncertainty in a mathematical model involving parameters which are not

known with certainty. By using first order analysis, the combined effect of uncertainty in a model formulation, as well as the use of uncertain parameters, can be assessed.

First order uncertainty analysis is characterized by two major components : (1) Single moment (variance) treatment of random variables ; and (2) the use of first order approximation of any functional relationship (e.g. the use of Taylor's series expansion). The first component implies that the random element of any variable is defined exclusively by its first nonzero moment or simply the varianceof the random variable itself. Thus the information pertaining to the character of a random variable is provided solely by its mean and variance.

The second component states that only the first order terms in a Taylor's series expansion will be utilized in the analysis of functional relationship containing random variables or processes. With exception of the evaluation of the mean (in which second order terms may be included for the purpose of accounting for correlation among variables), any attempt to retain terms higher than first order in the expansion requires more information about the random variables than those provided by their first and second moments (Cornell, 1972).

To present the general methodology of first order analysis, consider a random variable, Y, which is a function of N random variables (multivariate case). Mathematically, Y can be expressed as

$$Y = g(X) \qquad (1)$$

where $X = (X_1, X_2 \ldots\ldots, X_N)$, a vector containing N random variables X_i . Though the use of Taylor's expansion , the random variable Y can be approximated by

$$Y \overset{2}{=} g(\bar{X}) + \sum_{i=1}^{N} \left[\frac{\partial g}{\partial X_i} \right] (X_i - \bar{X}_i) + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{\partial^2 g}{\partial X_i \partial X_j} \right] (X_i - \bar{X}_i)(X_j - \bar{X}_j) \tag{2}$$

in which X = (X_1 , X_2 , ..., and X_N), a vector containing the mean of N random variables and = represents equal in the sense of a second - order approximation.

To consider the correlation among random variables X_s , the second order approximation of the mean (the expected value) of random variable Y is

$$\mu_Y = E[Y] \overset{2}{=} g(\bar{X}) + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{\partial^2 g}{\partial X_i \partial X_j} \right] \mathrm{cov}[X_i, X_j] \tag{3}$$

in which cov[X_i ,X_j] is the covariance between random variables X_i and X_j. It should be noted that the second term in equation (3) can be dropped if the random variables X_i are uncorrelated. In such a case, the resulting equation is the same as the first order approximation.

It follows that the first order approximation of the variable of Y is

$$\sigma_Y^2 = \mathrm{var}[Y] \overset{1}{=} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{\partial g}{\partial X_i} \right] \left[\frac{\partial g}{\partial X_j} \right] \mathrm{cov}[X_i, X_j] \tag{4}$$

If the X_i and X_j are uncorrelated , equation (4) reduces to

$$\sigma_Y^2 = \sum_{i=1}^{N} \left[\frac{\partial g}{\partial X} \right]_{\bar{X}}^2 \sigma_i^2 \qquad (5)$$

where = means equal in a first order sense and σ_i^2 = the variance corresponding to random variable X_i .

2.2 MONTE CARLO SIMULATION

MCS is a sampling procedure in which possible values for the input parameters are selected at random from appropriate pdfs and used in the model to produce estimates of the outputs. MCS is more complete than FOA and does not require assumptions of linearity and small parameters variances. It does require assumptions on the parameters pdfs and is computationally intensive generally requiring 1500 or more model runs. The approach illustrated in the following figure is actually quite simple in concept. After defining the probabilistic structure of the parameters, a parameter set is selected at random from the appropriate multivariate pdf. The model is run with these parameters and the output noted. This entire process is repeated many times (~1500) resulting in many (~1500) estimates for the outputs. These outputs are then analyzed probanilistically. The means and variances are calculated . Appropriate pdfs or probability plots are prepared and probabilistic statements made. Confidence intervals can be read directly from probability plots without making a distribution assumption about the outputs.

The importance of individual parameters in determining the uncertainty associated with particular variable can be assessed by computing the correlation between the model outputs and the input parameters. Those parameters that are highly correlated with the

model outputs are obviously important since changes in the values of these parameters will result in a corresponding change in the outputs from the models. On the other hand, if the correlations between outputs and parameters is low, the parameter is not very influential in determining the outputs.

Two checks that should be incorporated into any MCS are an examination of the correlation structure of the randomly generated parameter values and an examination of the pdfs of the randomly generated parameter values. The correlation structure of the parameters must statistically match the target correlation structure. If the intent was to generate independent rvs but in fact a high degree of correlation ends up in the generated parameters, the output variance will be incorrectly estimated. For example consider p1 and p2 as two parameters that are positively correlated with the model output but are uncorrelated with each other. If the generated "random" values of p1 and p2 are significantly positively correlated, the variance and the uncertainty in the output of the model will be exaggerated.

The fraction , F_i of the total variance in model output attributable to the ith parameter based on a MCS can be estimated by computing

$$F_i = \frac{r_{0,i}^2}{\sum_{i=1}^{p} r_{0,i}^2} \qquad (6)$$

where $r_{0,i}$ is the correlation between the output and the ith parameter and p is the number of uncertain parameters. This is a very rough approximation and may be used for guidance only. If the output variance is judged to be excessive, equation 8 and 9 can be used to determine which parameters are the biggest contributors to this uncertainty. Attempts can

then be made to reduce the variance on these influential parameters by getting better estimates for them.

Equation 7 can be used to estimate how much variance a particular parameter can have and still achieve a target variance on O. For example if V_t is the target variance, the target variance on parameter p_j can be estimated from

$$Var(p_j) = \frac{1}{S_j^2}\left[V_t - \sum_{\substack{i=1 \\ i \neq j}}^{P} S_i^2 Var(p_i)\right] \qquad (7)$$

The variance reduction calculations provide guidance only. They are not exact(obviously) since either FOA nor MCS is exact. Furthermore, it is not always possible to reduce the variance of the input parameters. The results of a MCS can be subjected to a multiple regression analysis in a further effort to identify the important input parameters. Regression coefficients relating a model output to the input parameters that are not statistically significant are likely of little importance in determining the model outputs.

2.3 PERFORMANCE EVALUATION OF THE DISTRIBUTIONS FOR DO

The idea of applying first order analysis for estimating the first two statistical moments of the DO deficit, along with a selection of an appropriate probability model for the DO deficit, straight forward and practical. However, among the various probability models that are commonly used, a practical question to be raised is, 'Which probability model (or models) best describes the random behaviour of the DO deficit in a stream?' In the following investigation, two parameteric i.e. normal and lognormal probability

distributions have been selected as candidates such that a wide spectrum of shapes are represented.

To evaluate the relative performance of each of the two candidate probability distributions considered, three performance criteria with respect to prediction ability are adopted herein: (1) Biasness (BIAS); (2) mean absolute error (MAE); and root mean squared error (RMSE). These criteria are used simultaneously in an attempt to identify the best probability model. These criteria are mathematically defined as

1. Biasness:

$$BIAS = \int_0^1 (x_{p,f} - x_p)dp \tag{8}$$

2. Mean absolute error:

$$MAE = \int_0^1 \left| x_{p,f} - x_p \right| dp \tag{9}$$

3. Root mean squared error:

$$RMSE = \left[\int_0^1 (x_{p,f} - x_p)^2 \right]^{0.5} \tag{10}$$

in which x_p = the true value of DO corresponding to the pth order of probability; and $x_{p,f}$ = the estimate of x_p determined from the assumed probability model, f, with its mean and variance estimated by the first-order analysis.

The test of goodness of fit is tested using the well known Kolmogorov-Smirnov test.

3.0 BASIC WATER QUALITY MODEL

Streeter and Phelps (1925) were among the first researchers to recognize the capacity of a water resource to receive and assimilate organic waste material depended on the oxygen economy. The first order reactions for deoxygenation and reaeration were combined to give the rate of change of oxygen deficit. The relationship among the parameters affecting the in-stream dissolved oxygen concentration is given by Equation (11):

$$D = \frac{K_d L_0}{K_a - K_d}(e^{-K_d t} - e^{-K_a t}) + D_0 e^{-K_a t} \qquad (11)$$

in which D = dissolved oxygen deficit (C_s - C) in milligrams per liter; C_s = dissolved oxygen saturation limit, in milligrams per liter; C = dissolved oxygen concentration, in milligrams per liter K_a = reareation rate coefficient (base e, per day); K_d = deoxygenation rate coefficient (base e, per day), L_0 = initial instream total ultimate biochemical oxygen demand, in milligrams per liter, D_0 = initial instream dissolved oxygen deficit, in mg/L.

and t = time of travel from D_0 to D (days).

3.1 Application of first-Order Analysis to obtain DO Deficit Profile:

First order uncertainty analysis of the basic Streeter Phelps Dissolved Oxygen Sag equation is presented in this section. Basically the first order uncertainty analysis will provide a measure of the uncertainty of the dependent variable, D in terms only of the uncertainty in independent variables: K_a , K_d, L_0, D_0 and t; i.e., percentage of the scatter of dissolved oxygen deficit predictions around the true deficit at any point along the sag curve can be assigned to each of the independent variables.

Taking the partial derivative of D with respect to each of the independent variables.

$$\frac{\partial D}{\partial D_0} = e^{-K_d t} \tag{12}$$

$$\frac{\partial D}{\partial L_0} = \frac{K_d}{K_a - K_d}(e^{-K_d t} - e^{-K_a t}) \tag{13}$$

$$\frac{\partial D}{\partial K_a} = \frac{-K_d L_0}{(K_a - K_d)^2}(e^{-K_d t} - e^{-K_a t}) + \frac{K_d L_0 t e^{-K_a t}}{(K_a - K_d)^2} - D_0 t e^{-K_a t} \tag{14}$$

$$\frac{\partial D}{\partial K_d} = \frac{K_d L_0}{(K_a - K_d)^2}(e^{-K_d t} - e^{-K_a t}) - \frac{K_d L_0 t e^{-K_d t}}{K_a - K_d} \tag{15}$$

$$\frac{\partial D}{\partial t} = \frac{K_d L_0}{K_a - K_d}(K_a e^{-K_d t} - K_d e^{-K_a t}) - K_a D_0 e^{-K_a t} \tag{16}$$

The first -order approximation to the total uncertainty in the dissolved oxygen deficit is obtained by applying Equ. (5), the resulting equation is given by Equ. (17).

$$\sigma_D = \left(\sum_{i=1}^{5} |C_i|^2\right)^{\frac{1}{2}} \tag{17}$$

The terms of Eq. 17 are defined by Eq. 18

$$C_1 = \frac{\partial D}{\partial D_0}\sigma_{D_0}; C_2 = \frac{\partial D}{\partial L_0}\sigma_{L_0}; C_3 = \frac{\partial D}{\partial K_a}\sigma_{K_a}$$
$$C_4 = \frac{\partial D}{\partial K_d}\sigma_{K_d}; C_5 = \frac{\partial D}{\partial t}\sigma_t \qquad (18)$$

For Eqs. 17 and 18, the symbol S represents the standard deviation of particular variable. Thus, Eq. 17 shows that each of the independent variables contributes to the dispersion of D in a manner proportional to its own variance, S^2, and proportional to a factor $\left|(\partial D/\partial)\right|^2_{mean}$ which is related to the sensitivity of changes in D to changes in the independent variable (Benjamin and Cornell, 1970).

Application of the method of first order uncertainty analysis to the BOD-DO system requires estimates of mean parameter values and standard deviations of D_O, L_O, K_d, K_a, and t.

3.2 Application of First-Order Analysis to obtain Critical DO Sag:

The lowest or critical point of DO curve is important as it gives the greatest deficit in dissolved oxygen. The critical time can be obtained by differentiating the sag curve equation (11) with respect to time and placing the resulting expression equal to zero. The point of the minimum oxygen content, in t ms of time, t_c, is thus obtained as

$$t_c = \frac{1}{k_d(f-1)} \ln\left[f\left(1+\frac{D_0}{L_0}\right) - f^2 \frac{D_0}{L_0} \right] \qquad (19)$$

It is clear from Ewu. (19) that t_c depends upon four independent variables namely k_d, k_e, L_0, and D_0. Taking the partial derivative of t_c with respect to each of the independent variables.

$$\frac{\partial t_c}{\partial k_d} = \frac{1}{k_d^2(f-1)^2} \ln\left[f\left(1+\frac{D_0}{L_0}\right) - f\frac{D_0}{L_0} \right] + \frac{1}{k_d^2(f-1)\left[1+\frac{D_0}{L_0}(1-f)\right]}\left(\frac{2fD_0}{L_0} - \frac{D_0}{L_0} - 1\right) \qquad (20)$$

$$\frac{\partial t_c}{\partial D_0} = -\frac{1}{k_d L_0\left[1-(f-1)\frac{D_0}{L_0}\right]} \qquad (21)$$

$$\frac{\partial t_c}{\partial k_e} = -\frac{1}{k_d^2(f-1)^2} \ln\left[f\left(1+\frac{D_0}{L_0}\right) - f\frac{D_0}{L_0} \right] + \frac{1}{k_d k_e(f-1)\left[1+\frac{D_0}{L_0}(1-f)\right]}\left(1+\frac{D_0}{L_0} - \frac{2fD_0}{L_0}\right) \qquad (22)$$

$$\frac{\partial t_c}{\partial L_0} = \frac{D_0}{k_d L_0^2\left[1-(f-1)\frac{D_0}{L_0}\right]} \qquad (23)$$

The mean value of t_c is obtained by substituting the mean values of k_d, k_e, L_0, and D_0

in Equ. (19). The equation for variance in t_c is obtained by substituting various senstivity coefficients in Equ. (5), as given below

$$\sigma_{t_c}^2 = \left(\frac{\partial t_c}{\partial k_d}\right)^2 \sigma_{k_d}^2 + \left(\frac{\partial t_c}{\partial k_a}\right)^2 \sigma_{k_a}^2 + \left(\frac{\partial t_c}{\partial L_0}\right)^2 \sigma_{L_0}^2 + \left(\frac{\partial t_c}{\partial D_0}\right)^2 \sigma_{D_0}^2 \qquad (24)$$

The maximum DO deficit D_c is obtained by substituting $t = t_c$ in Equ. (11) from Equ. (19), the following expression is evolved

$$D_c = \frac{k_d L_0}{k_a} e^{-k_d t_c} \qquad (25)$$

It is seen from Equ. (24) that D_c is a function of four independent variables namely k_d, k_a, L_0, and t_c. Taking the partial derivative of D_c with respect to each of the independent variables to get the respective senstive coefficient corresponding to each of the independent variable.

$$\frac{\partial D_c}{\partial k_d} = \frac{L_0(1 - k_d t_c)}{k_a} e^{-k_d t_c} \qquad (26)$$

$$\frac{\partial D_c}{\partial k_a} = -\frac{k_d L_0}{k_a^2} e^{-k_d t_c} \qquad (27)$$

$$\frac{\partial D_c}{\partial L_0} = \frac{k_d}{k_a} e^{-k_d t_c} \tag{28}$$

$$\frac{\partial D_c}{\partial t_c} = -\frac{k_d^2 L_0}{k_a} e^{-k_d t_c} \tag{29}$$

The variance in D_c is calculated using the following equation

$$\sigma_{t_c}^2 = \left(\frac{\partial t_c}{\partial k_d}\right)^2 \sigma_{k_d}^2 + \left(\frac{\partial t_c}{\partial k_a}\right)^2 \sigma_{k_a}^2 + \left(\frac{\partial t_c}{\partial L_0}\right)^2 \sigma_{L_0}^2 + \left(\frac{\partial t_c}{\partial D_0}\right)^2 \sigma_{D_0}^2 \tag{30}$$

4.0 DATA SELECTION

The independent variables of Equ. 11 are subject to wide variations from stream to stream or even from reach to reach of the same stream. Global or even regional mean values and standard deviations would be difficult to estimate and would probably not permit useful conclusions to be drawn. Four classes of streams, defined by the self - purification ratio, $f = K_a/K_d$, were described by Fair et al. (1968) as given in table 1.

TABLE 1 - Ranges of Data by Stream class

Stream-class description	f (dimensionless)	K_a (per day)	K_d (per day)	V (meter/sec.)	H (meters)
Sluggish	1.25-1.50	0.05-0.10	0.033-0.08	0.03-0.15	3.05-6.10
Low-velocity	1.50-2.00	0.10-1.00	0.050-0.67	0.03-0.15	0.92-3.05
Moderate-velocity	2.00-3.00	1.00-5.00	0.500-2.50	0.15-0.61	0.61-1.52
Swift	3.00-5.00	1.00-10.0	0.200-3.33	o.61-1.83	0.61-3.05

In the present study following data given in table 2 is used in the uncertainty analysis. This set of data is taken from Burges and Lettenmaier (1975).

Table 2: Parameters Used in Uncertainty Analysis

Parameters	Mean	Standard Deviation	Coefficient of variation
Initial BOD, L_0 (ppm)	12.15	1.00	0.08
Initial DO Deficit, D_0 (ppm)	1.00	0.50	0.50
Deoxygenation Coefficient k_d (per day)	0.331	0.10	0.32
Reoareation Coefficient k_e (per day)	0.690	0.20	0.29
Travel Time, t, (days)	—	—	0.25

For a detailed study, the uncertainty analysis could be applied to individual streams as given in table 1. Furthermore, a wide range of parameter values and combinations could be selected to permit study of the relative importance of the variables for variety of stream conditions.

5.0 Results and Discussions

5.1 Estimation of DO Profile

The mean DO profiles alongwith their standard deviation obtained by first-order uncertainty analysis and Monte-Carlo simulations (assuming input to be Normally distributed) are shown in Fig.1 and Fig. 2 respectively. Comparisons between Monte Carlo simulations and first order analysis are also shown in Table-3 and Fig. 3.

The observations indicate that the estimated mean value of dissolved oxygen (DO) appreciably different for either method of estimation. It is clear from Fig. 3 that DO levels are more or less same by both the methods up to the point of maximum sag. After this point,both the profiles depart from each other and the deviation between them increases with distance and saturates after a certain distance and then remain more or less constant. Furthermore, it is observed that the DO levels given by first order analysis are always on the higher side in the post sag region of the stream.

The standard deviation estimated by both methods indicates that the first order approach is quite satisfactory as it has less variance than the Monte-Carlo simulations. The standard deviation increases to a maximum and diminishes in magnitude with distance along the stream. The point of maximum uncertainty is somewhat down stream of the minimum dissolved oxygen level.

Now, assuming all the parameters of DO model are to be lognormally distributed. The resulting simulated profile is shown in Fig. 4. The comparison between the first order and Monte-Carlo simulations (assuming all the input parameter to be log normally distributed) is shown in fig. 5 and table 4.

Table 3: Comparison of First-Order Uncertainty Estimates and Monte Carlo Estimates of DO (ppm) under Normal Distribution

Time of Travel (days)	First Order Analysis		Monte Carlo Analysis		Skewness
	Mean	Std.Deviation	Mean	Std. Deviation	
0.0	8.00	0.50	7.93	0.73	-6.4
0.5	6.73	0.63	6.70	0.80	-2.99
1.0	6.07	0.79	6.09	0.90	-1.58
1.5	5.81	0.85	5.83	0.97	-1.32
2.0	5.79	0.89	5.81	1.00	-1.28
2.5	5.92	0.91	5.94	1.03	-1.39
3.0	6.14	0.93	6.10	1.07	-1.60
3.5	6.39	0.95	6.33	1.10	-1.70
4.0	6.67	0.95	6.55	1.13	-1.88
4.5	6.93	0.94	6.77	1.14	-1.93
5.0	7.18	0.92	6.99	1.13	-2.24
5.5	7.42	0.88	7.22	1.13	-2.46
6.0	7.63	0.84	7.37	1.13	-2.67
6.5	7.81	0.79	7.51	1.13	-2.76
7.0	7.98	0.73	7.70	1.06	-3.37
7.5	8.12	0.68	7.81	1.06	-3.31
8.0	8.25	0.62	7.91	1.06	-3.52
8.5	8.36	0.57	8.04	1.03	-4.03
9.0	8.45	0.52	8.12	1.01	-4.21
9.5	8.53	0.47	8.23	0.96	-4.98
10.0	8.60	0.42	8.27	0.99	-4.86
10.5	8.66	0.38	8.34	0.95	-5.17
11.0	8.71	0.34	8.42	0.88	-5.80
11.5	8.75	0.30	8.47	0.86	-6.04
12.0	8.79	0.27	8.51	0.86	-6.33

Table 4: Comparison of First-Order Uncertainty Estimates and Monte Carlo Estimates of DO (ppm) under Log Normal Distribution

Time of Travel (days)	First Order Analysis		Monte Carlo Analysis		Skewness
	Mean	Std. Dev.	Mean	Std. Dev.	
0.0	8.00	0.50	8.02	0.50	-1.54
0.5	6.73	0.63	6.78	0.64	-0.83
1.0	6.07	0.79	6.15	0.77	-0.56
1.5	5.81	0.85	5.91	0.82	-0.58
2.0	5.79	0.89	5.88	0.84	-0.56
2.5	5.92	0.91	6.00	0.84	-0.48
3.0	6.14	0.93	6.18	0.86	-0.53
3.5	6.39	0.95	6.39	0.89	-0.55
4.0	6.67	0.95	6.64	0.87	-0.54
4.5	6.93	0.94	6.83	0.87	-0.57
5.0	7.18	0.92	7.07	0.85	-0.73
5.5	7.42	0.88	7.29	0.83	-0.86
6.0	7.63	0.84	7.45	0.84	-1.07
6.5	7.81	0.79	7.63	0.78	-1.30
7.0	7.98	0.73	7.71	0.79	-1.16
7.5	8.12	0.68	7.91	0.75	-3.31
8.0	8.25	0.62	8.00	0.72	-3.52
8.5	8.36	0.57	8.13	0.66	-4.03
9.0	8.45	0.52	8.20	0.68	-4.21
9.5	8.53	0.47	8.28	0.61	-4.98
10.0	8.60	0.42	8.38	0.59	-4.86
10.5	8.66	0.38	8.43	0.56	-5.17
11.0	8.71	0.34	8.49	0.53	-5.80
11.5	8.75	0.30	8.56	0.48	-6.04
12.0	8.79	0.27	8.59	0.43	-6.33

It is very interesting to note that in the pre-sag region the first order profile over estimates the level of DO sag, while in the post-sag region, the Monte-Carlo profile gives higher estimates for DO levels in the stream. Because the pre-sag region is more important from the management of water quality point of view the over-estimation of DO sag is on the safer side. Therefore, first order analysis is more suitable than the Monte-Carlo simulations. Furthermore, the first order profile has more uncertainty in the pre-sag than the Monte-Carlo profile. But at the point of critical sag the difference in the standard deviation determined by first order and Monte-Carlo simulations is of the order of 0.05 ppm which can be neglected for the sake of simplicity. This difference increases after the sag point and then reduces along the stream and at a certain point they become equal and then standard deviation determined by Monte-Carlo simulations over takes the standard deviation given by first order. This difference first increases and then remains constant after getting saturated in the post sag region.

5.2 Estimation of Location of minimum DO and minimum DO distribution.

The mean and standard deviation for minimum DO level and its location on the river determined by first order analysis and Monte-Carlo simulations assuming all the parameters are normally distributed have been given in table-5. The probability distribution functions for minimum DO levels determined by first order analysis and Monte-Carlo simulations using normally distributed input are shown in fig. 6 to fig.13 respectively along with their cumulative distribution functions. It is observed that both the mean value and standard deviation of DO determined by first order analysis are on higher side than the corresponding values obtained by Monte-Carlo simulations.

Furthermore, it is noticed that the DO distribution determined by Monte-Carlo simulations is skewed towards left indicting the occurrence of lower DO levels in the stream at the critical location which is of immense importance for the management of water quality in a stream. This particular phenomenon is eclipsed by the first-order analysis.

Now, assuming all the input parameters of water quality model to be lognormally distributed, the minimum DO distribution is determined as shown in fig. 10 and fig.11. The comparison of this distribution with that of first order analysis is given in table 6. It is again observed that the mean and standard deviation given by first order analysis are on higher side than those of the Monte-Carlo simulation. Comparing table 5 and table 6, the influence of distribution type of input parameters on the minimum DO is clearly visible. The assumption of lognormal distribution for input parameters reduces the uncertainty in DO level, while increases its magnitude of mean value which is more or less equal to that of the mean value given by first-order analysis. For further insight into the DO distribution at the critical section of stream, the probability density function along with its cumulative distribution is plotted in fig. 12 and fig. 13, assuming log-normal distribution for output DO with mean and standard deviation obtained by first order analysis. It is noticed that this distribution is more or less same as the DO distribution obtained by Monte-Carlo simulation assuming all the input parameters to be log normally distributed.

Now question arises which distribution is the best to predict the DO distribution at the critical section. For answering this question two criteria have been used.

Table 5: Critical DO in the river reach under Normal Distribution

Method of Analysis	Dissolved Oxygen (ppm)		Travel Time (days)	
	Mean	Standard Dev.	Mean	Standard Dev.
First-Order Analysis	5.770	1.062	1.785	0.138
Monte-Carlo Simulation	5.663	1.001	1.864	0.589

Table 6: Critical DO in the river reach under Log Normal Distribution

Method of Analysis	Dissolved Oxygen (ppm)		Travel Time (days)	
	Mean	Standard Dev.	Mean	Standard Dev.
First-Order Analysis	5.770	1.062	1.785	0.138
Monte-Carlo Simulation	5.705	0.834	1.858	0.443

5.3 Comparison of Statistics from K-S Test:

The analysis of the goodness of fit performance criteria was conducted at the critical location using K-S test as given below:

5.3.1 Assuming Normally Distributed Input:

Table 7: K-S Test for DO, assuming DO to be normally distributed

Method of Analysis	Max. Diff.	Probability (2-Tail)
Monte Carlo (Output is normally distributed)	0.068	0.00
Monte Carlo (input is log-normally distributed)	0.100	0.00
First-Order Analysis	0.016	0.964
First-Order Analysis (Output is lognormally distributed)	0.048	0.018

5.3.2 Assuming Log Normally Distributed Input:

Table 8: K-S Test for DO, assuming DO to be lognormally distributed

Method of Analysis	Max. Diff.	Probability (2-Tail)
Monte Carlo (Output is normally distributed)	0.041	0.068
Monte Carlo (input is log-normally distributed)	0.072	0.00
First-Order Analysis	0.016	0.964
First-Order Analysis (Output is lognormally distributed)	0.052	0.009

From table 7 and table 8, it could be concluded that the distribution of output DO can not be assumed to be lognormally distributed. Because assuming it to be log-normally distributed has increased the Max. Diff. from 0.068 to 0.101 when input is

assumed to be normally distributed and from 0.041 to 0.072 when input is assumed to be lognormally distributed. Similar trend is also observed with the first order analysis. The influence of type of input variable distributions could be seen by comparing Max Diff. statistics for Monte-Carlo simulations given in table 7 and table 8. The assumption of lognormal distribution for input variables improves the Max diff. statistics from 0.068 to 0.041. This indicates that it is better to assume log normal distribution for the input parameters instead of normal distribution.

5.4 Prediction Performance Evaluation Criteria:

The analysis of the prediction ability criteria was also conducted at the critical location to see which distribution is better for prediction purposes.

Table - 9: Test for the predictability criteria

Criteria	True Distr.:Monte Carlo simulatin with normally distributed input Assumed Distr.: Normal distribution with first-order parameters	True Distr.:Monte Carlo simulatin with log- normally distributed input Assumed Distr.: Normal distribution with first-order parameters
BIAS	0.293	0.109
MAE	0.382	0.436
RMSE	0.471	0.319

It is clear from table 9 that the assumption of lognormal distribution for input parameters improves the Bias and RMSE confirming the above conclusion that input parameters should be assumed to be lognormally distributed instead of normally distributed.

5.5 Uncertainty in the Location of Point of Minimum DO on the Stream:

The location of the critical DO in a stream is also not certain at a fixed point. It varies with the variation in the input variables. The uncertainty analysis of the location of critical DO is summarized in the following table (Table-10).

Table 10: Variation in critical time obtained using different methods

Method of Analysis	Travel Time (days)	
	Mean Value	Standard Dev.
First-Order Analysis	1.785	0.138
Monte Carlo with normally distributed input	1.864	0.589
Monte Carlo with log-normally disttributed input	1.858	0.443

From table 10, it is seen that the first order analysis under estimates both the mean value and standard deviation of the critical travel time. Whereas, the mean travel time estimated by the Monte-Carlo Simulations is more or less same irrespective of the types of input variable distributions. But, they too differ with respect to the uncertainty associated with the travel time considerably. As the earlier discussion with respect to the minimum DO favors the lognormal distribution for the input variables, it is justified that the standard deviation of 0.443 in travel time given by the Monte Carlo simulations, assuming the input to be lognormally distributed is reasonably acceptable.

6.0 Conclusions

6.1 Critical DO & its Location

1. Neither normal distribution nor lognormal distribution is found justified to represent the output DO distribution. However, if one has to make a choice between them, normal distribution is found to be preferred.

2. Lognormal distribution is found to be preferred distributions for the input variables.

3. For estimation of travel time, Monte-Carlo simulation with lognormal input variables is found to be a preferred method.

4. The mean value of minimum DO at the critical location is found to be more or less same irrespective of the method used. But level of uncertainty associated with the minimum DO is found to be considerably different. The Monte-Carlo simulations with log normal distribution input is found to give the least uncertainty in the level of minimum DO levels.

6.2 DO Profile Along the Stream

1. The magnitude of DO level determined by first order analysis and Monte Carlo simulations using normally distributed input is almost equal in the pre-sag region. The Monte-Carlo simulation using lognormally distributed input gives higher DO levels in the pre-sag region. Therefore, the first-order analysis being on the safer side for water quality management activities, is found to be more justified.

2. From the uncertainty point of view, first order analysis having less standard deviation than the Monte Carlo simulation, is found to be more suitable to predict the uncertainty in the DO profile.

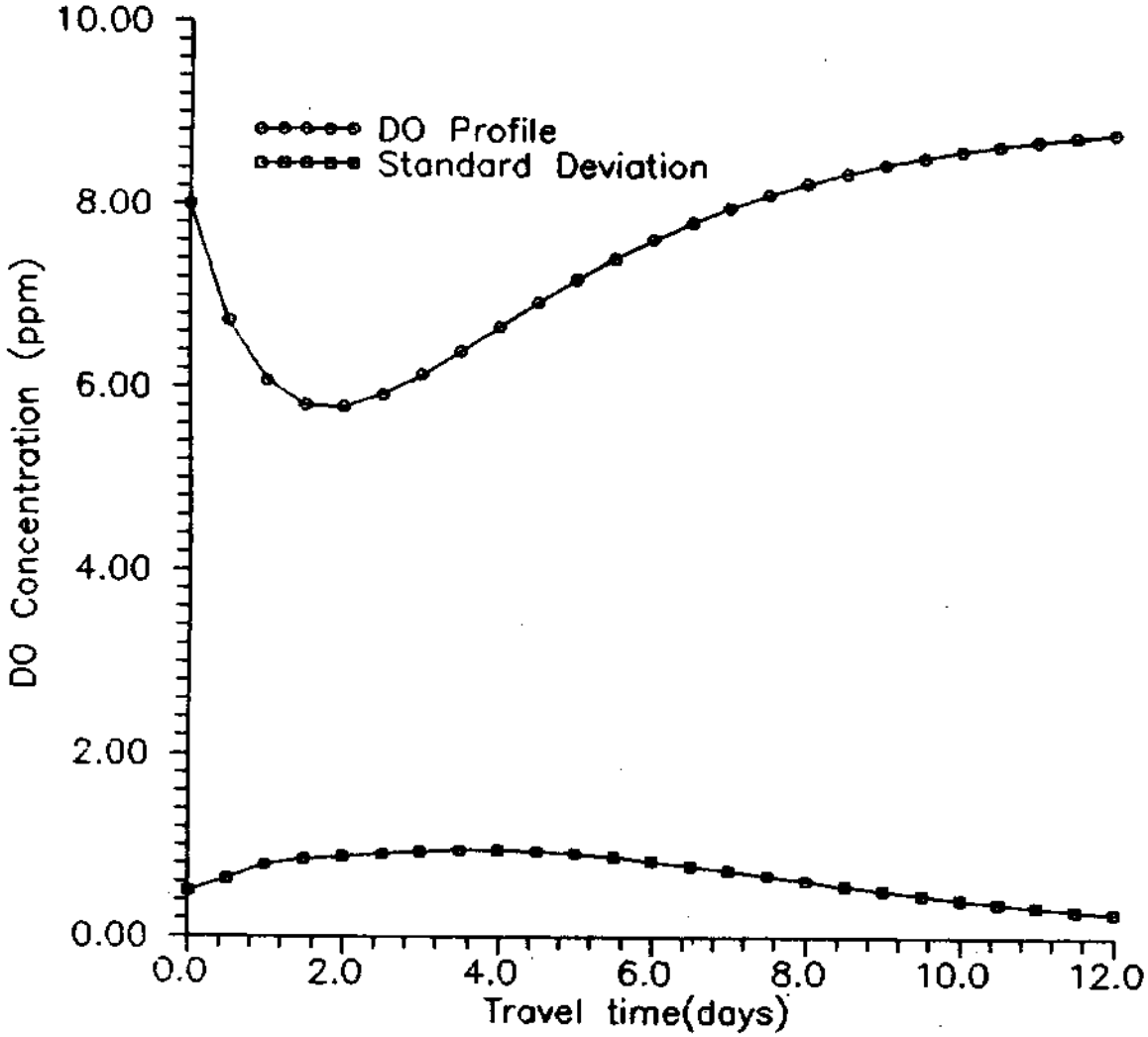

Fig.1 DO Profile using First Order Analysis

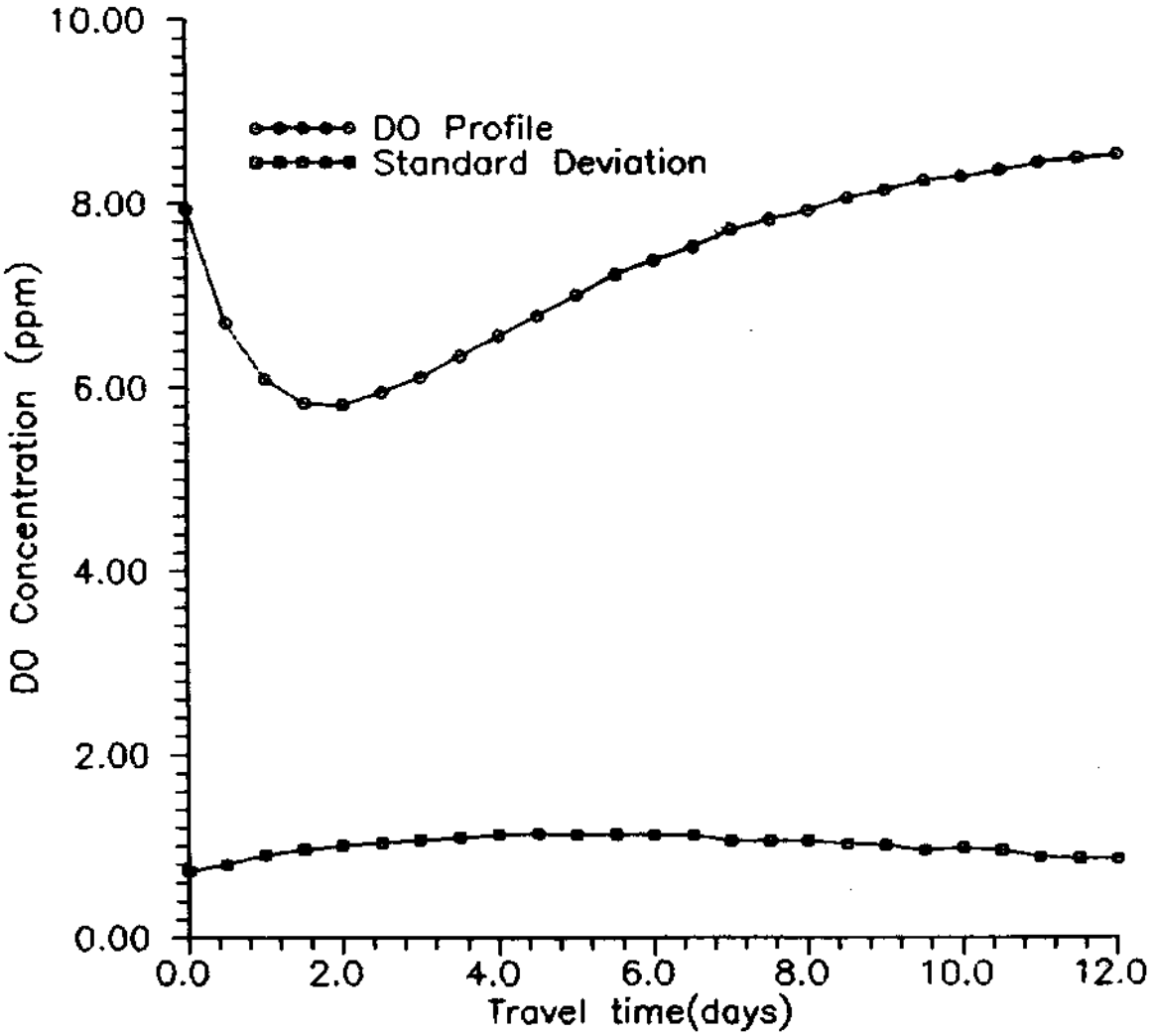

Fig. 2 DO Profile using Monte Carlo Simulation

2.

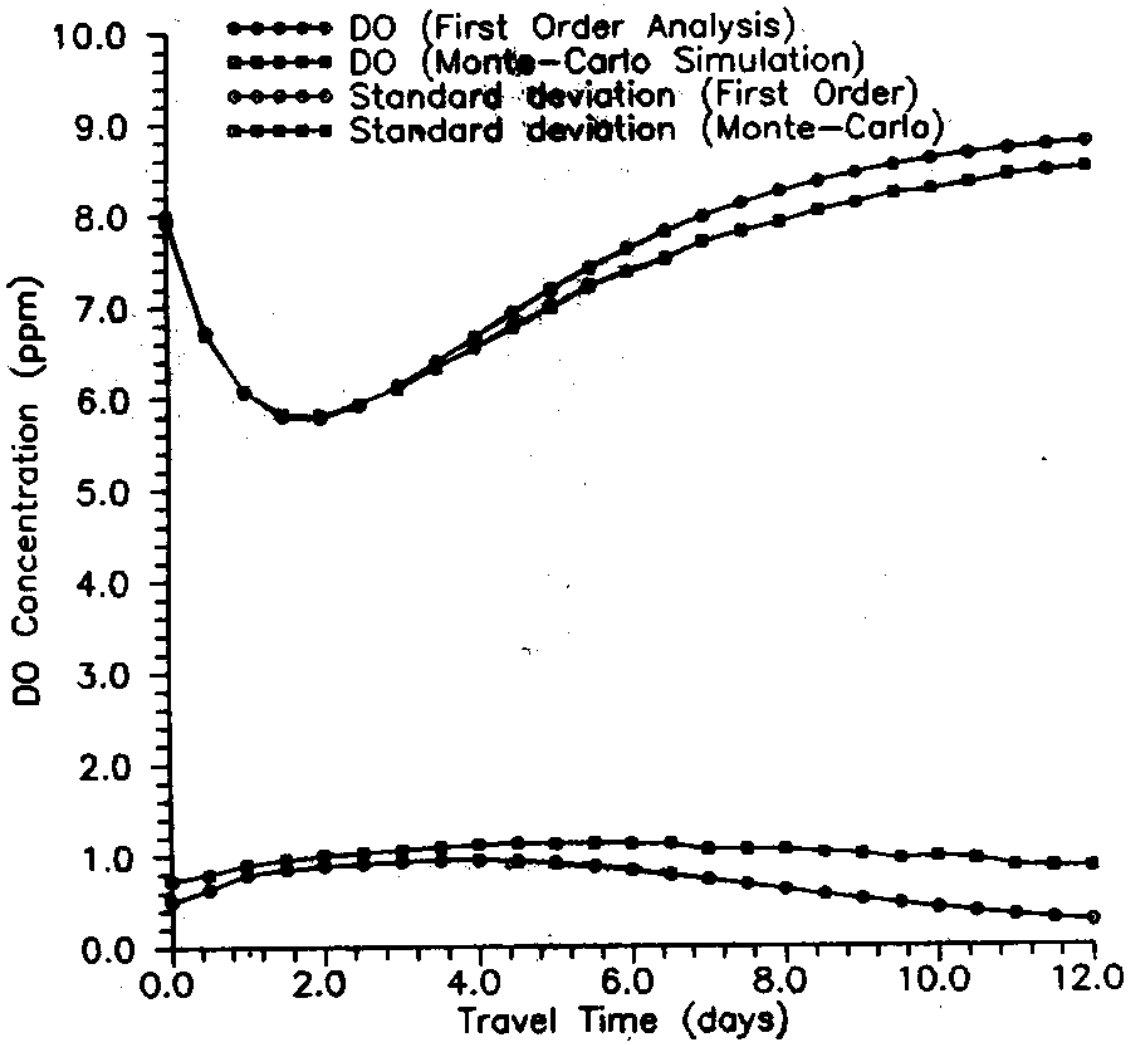

Fig. 3 Comparison of Monte Carlo Simulation and First order Analysis

3

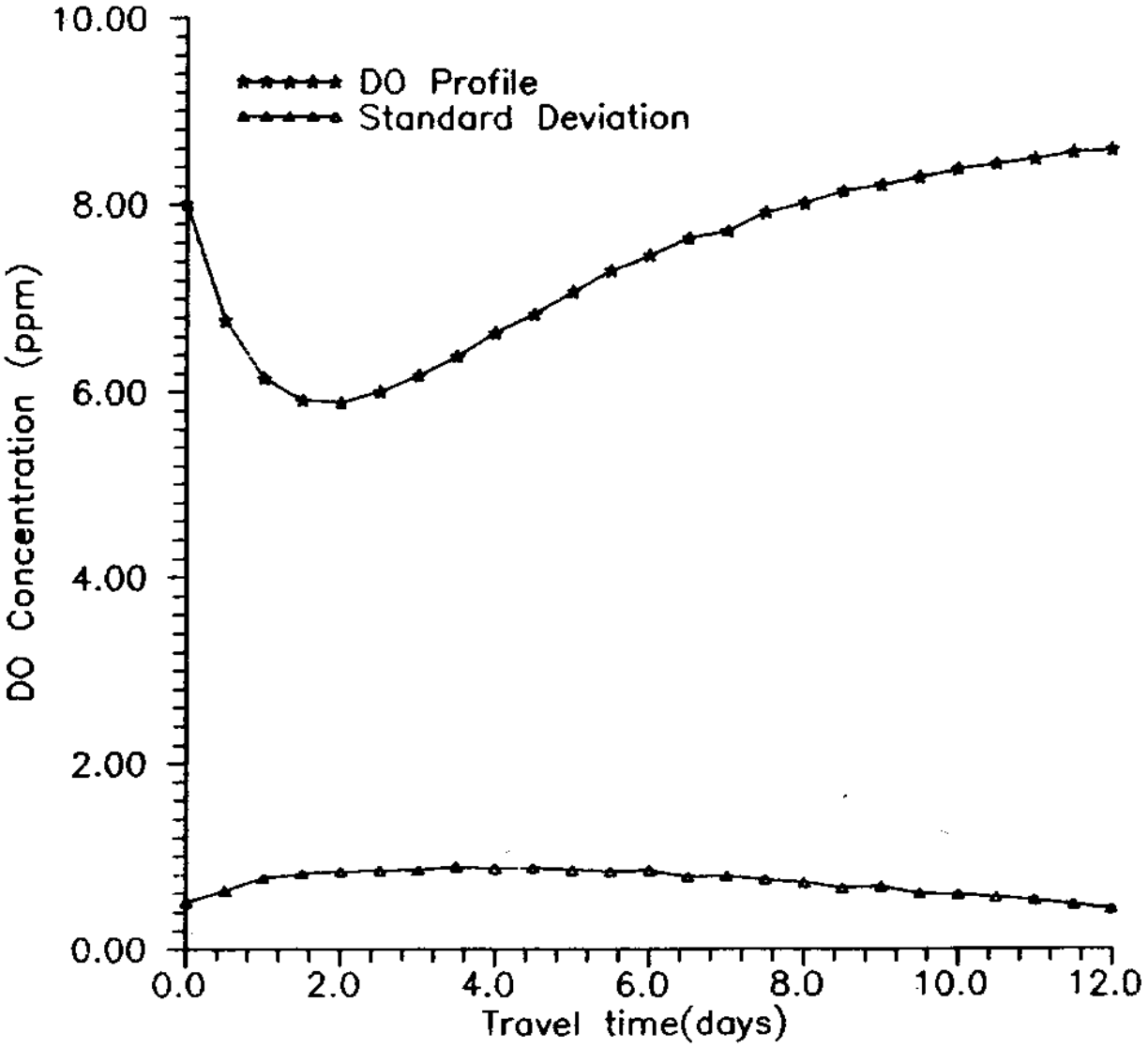

Fig. 4 DO Profile using Monte Carlo Simulation assuming all the input variables are Log–Normally Distributed.

4

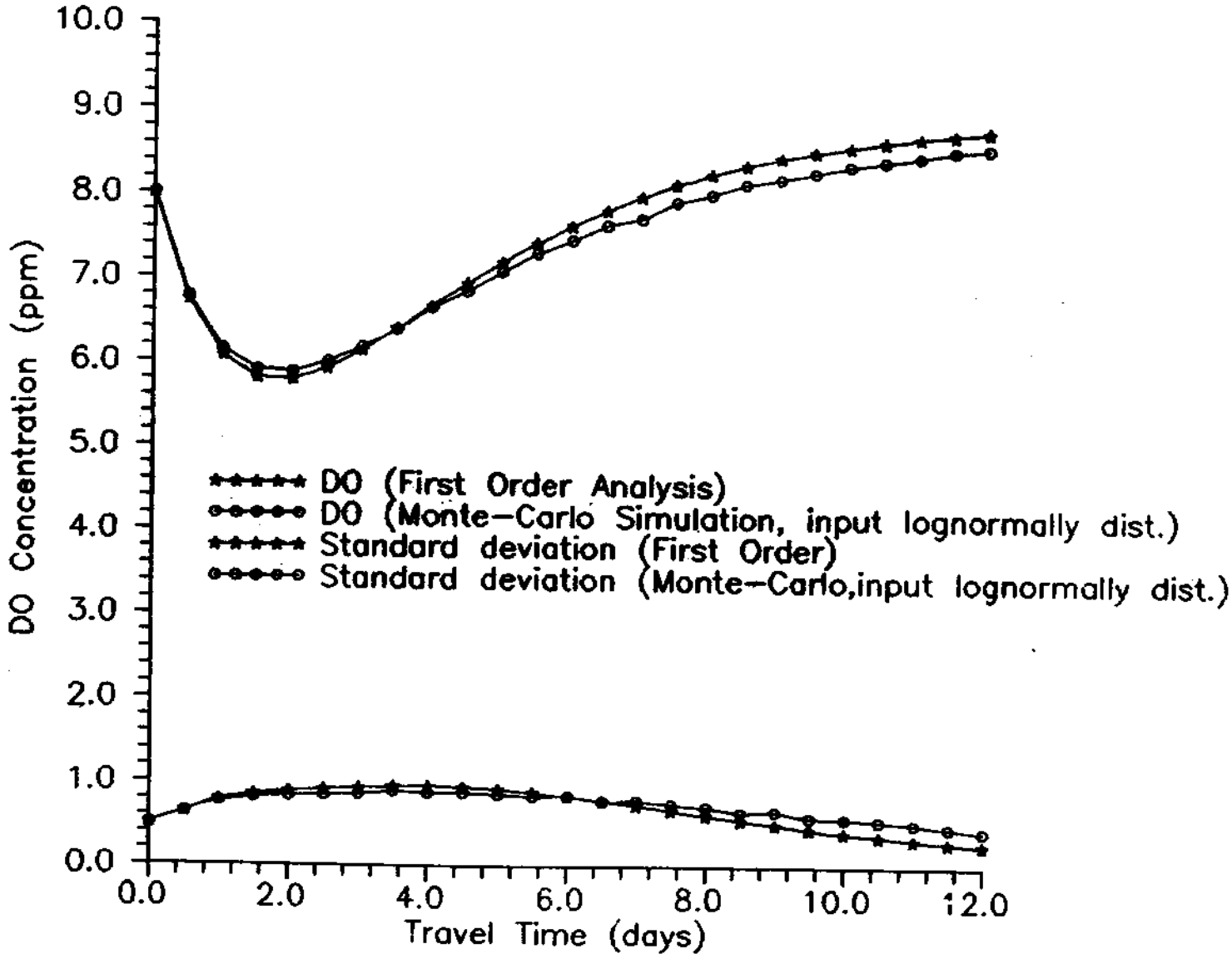

Fig.5 Comparison of Monte Carlo Simulation and First order Analysis

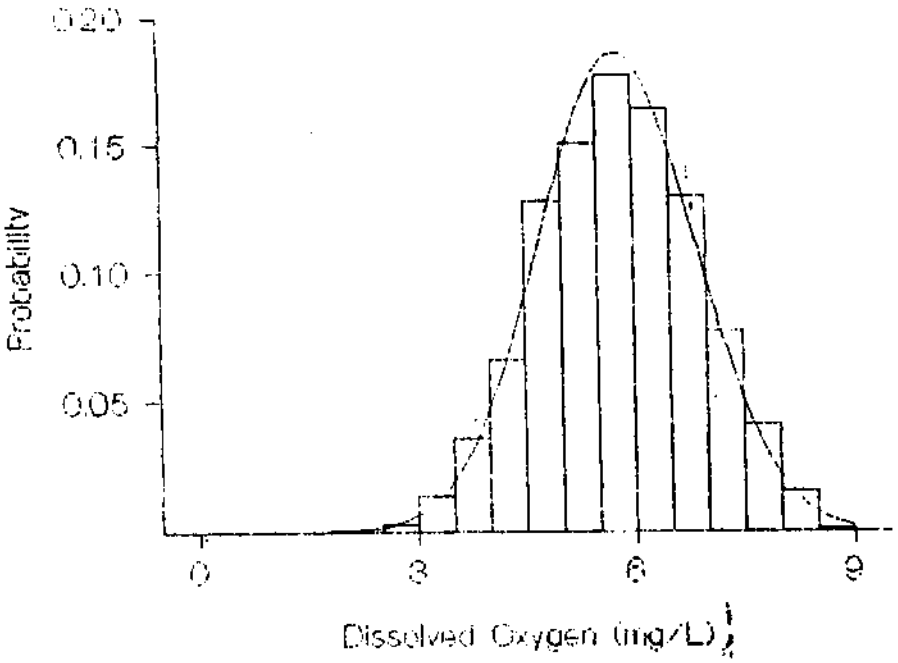

Fig. 6 Probability density function for critical DO assuming it to be Normally distributed with its parameters determined by First-Order Uncertainty Analysis

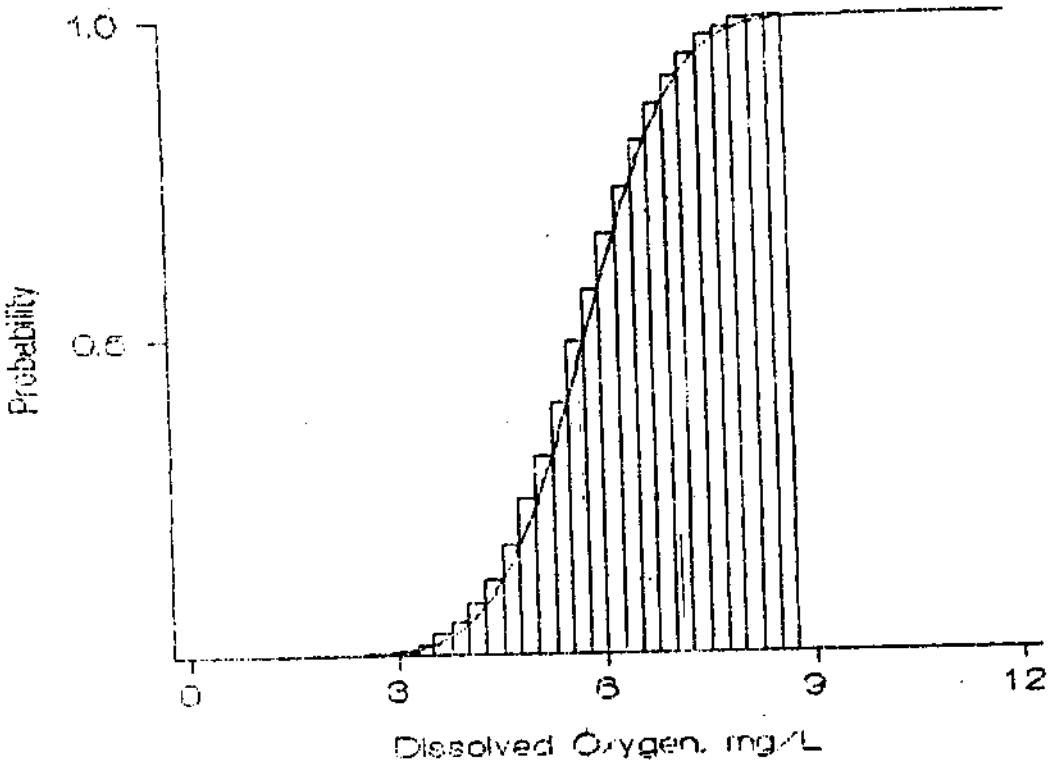

Fig. 7 CPDF for DO Using First-Order Analysis and assuming Normal Distribution

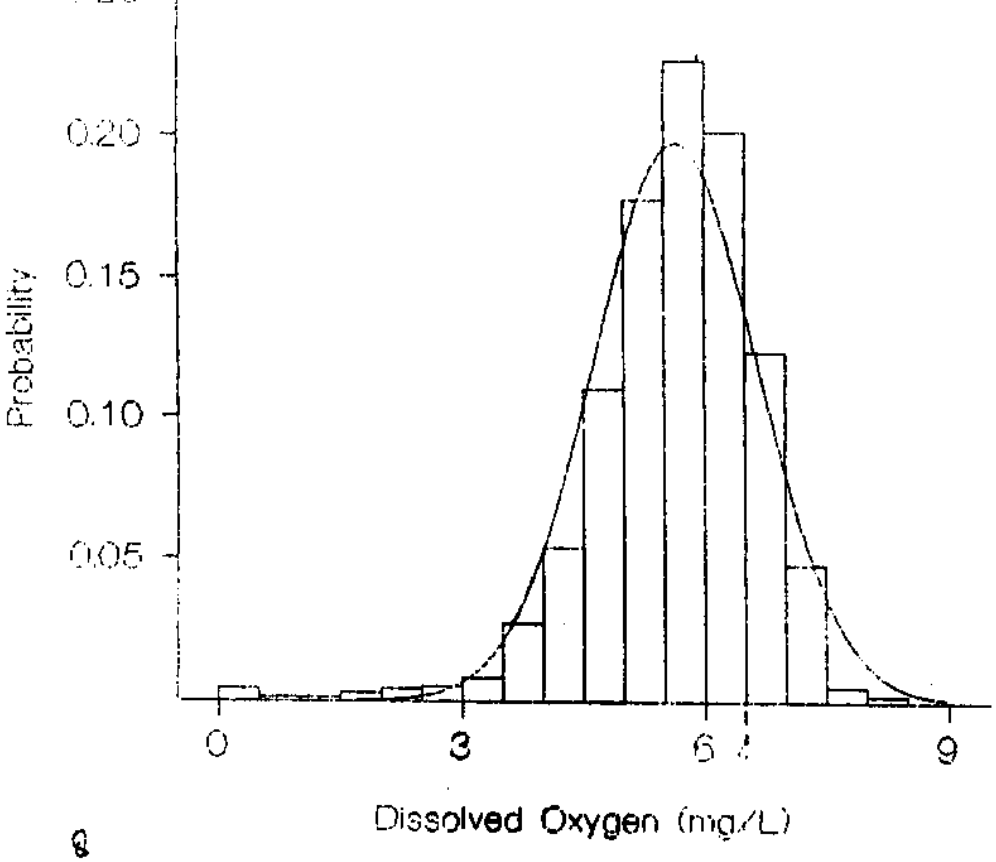

Fig. 8 Probability density function for critical DO, determined by Monte-Carlo analysis assuming all the input parameters to be Normally distributed.

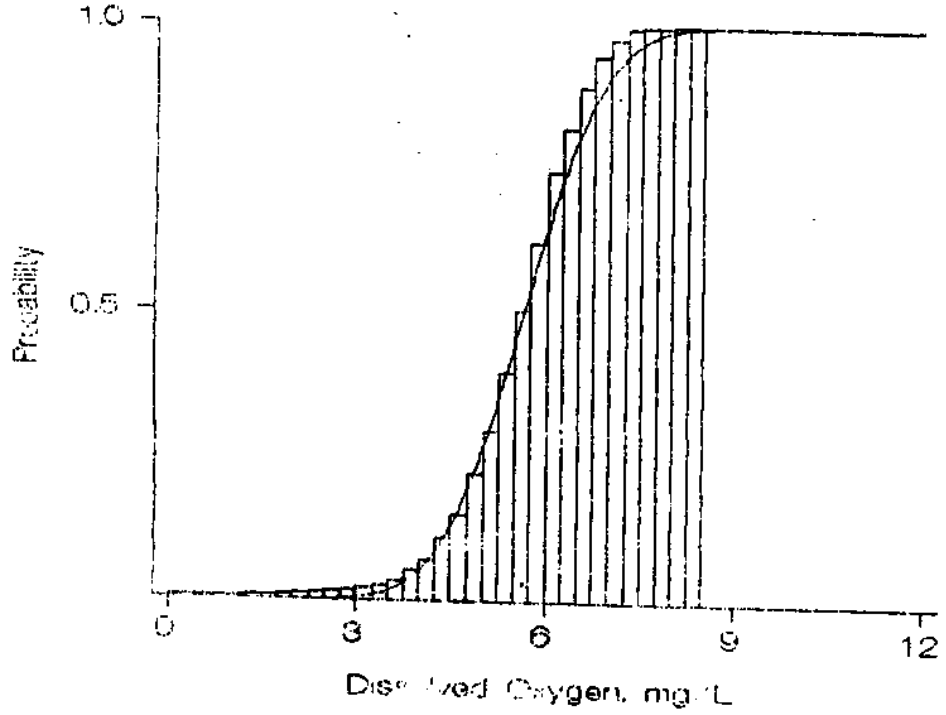

Fig 9 CPDF for DO using Monte-Carlo Analysis and Normally Distributed Input

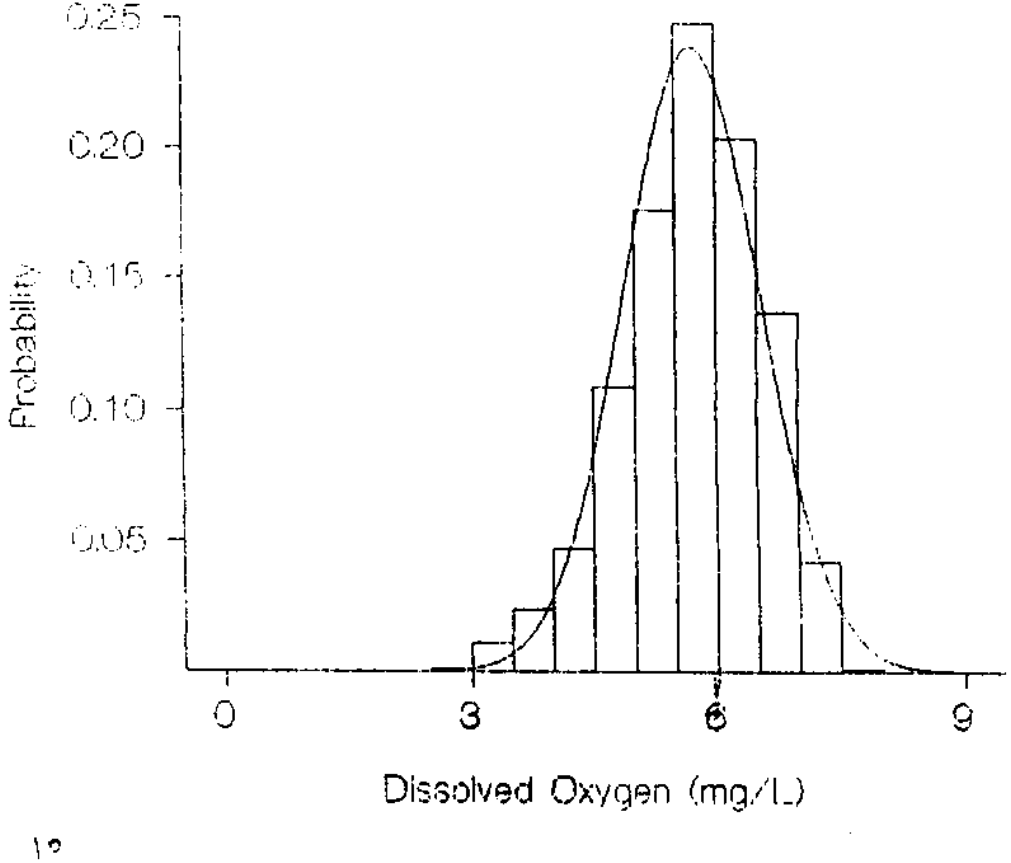

Fig.10 Probability density function for critical DO, determined by Monte-Carlo analysis assiming input parameters to be Log-Normally distributed.

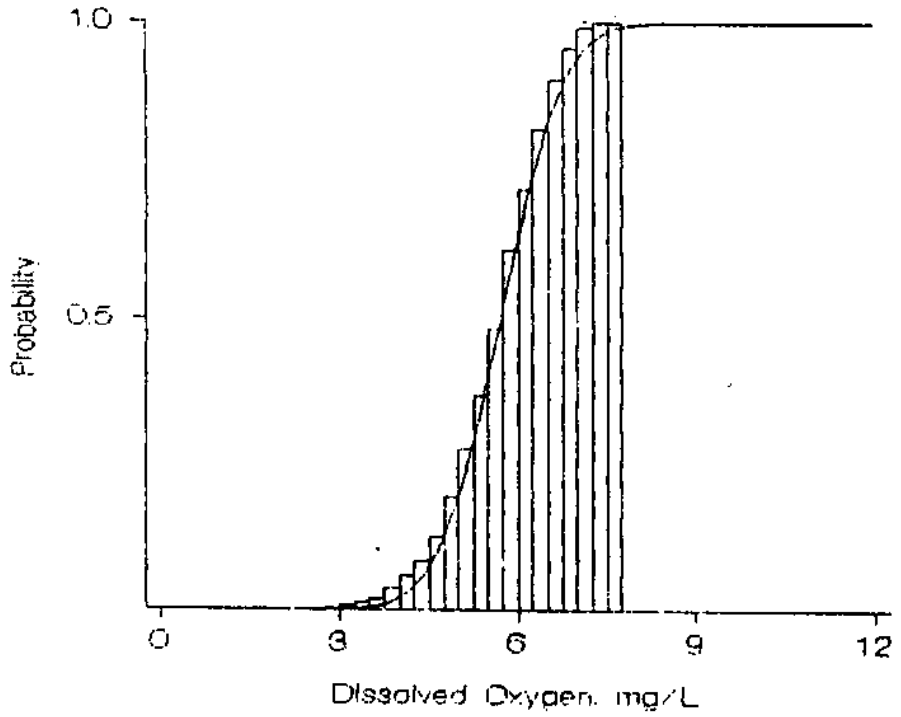

Fig.11 CPDF of DO Using Monte-Carlo Analysis and Log Normal Distributed Input Parameters

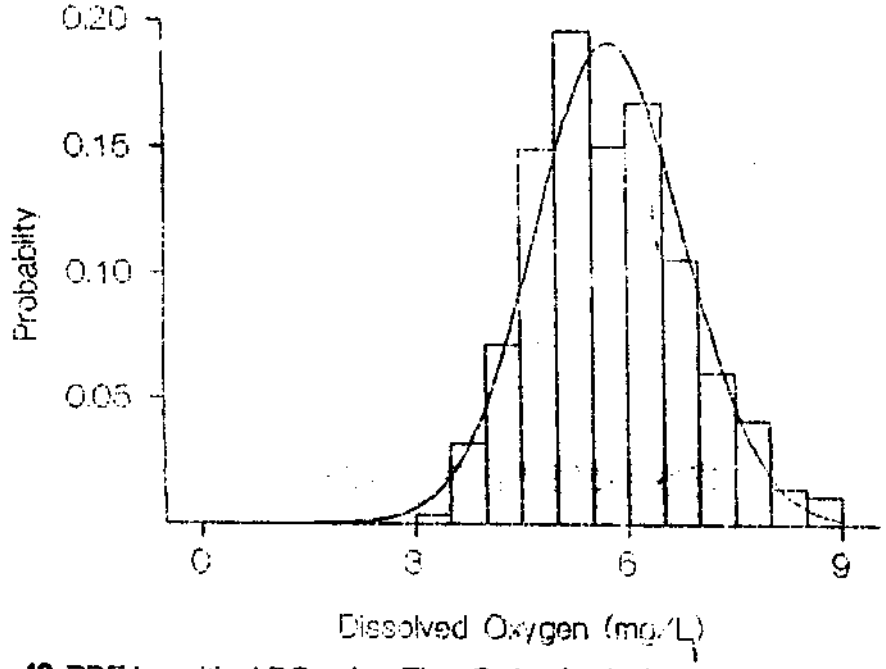

Fig.12 PDF for critical DO using First-Order Analysis and Log-Normal Distr.

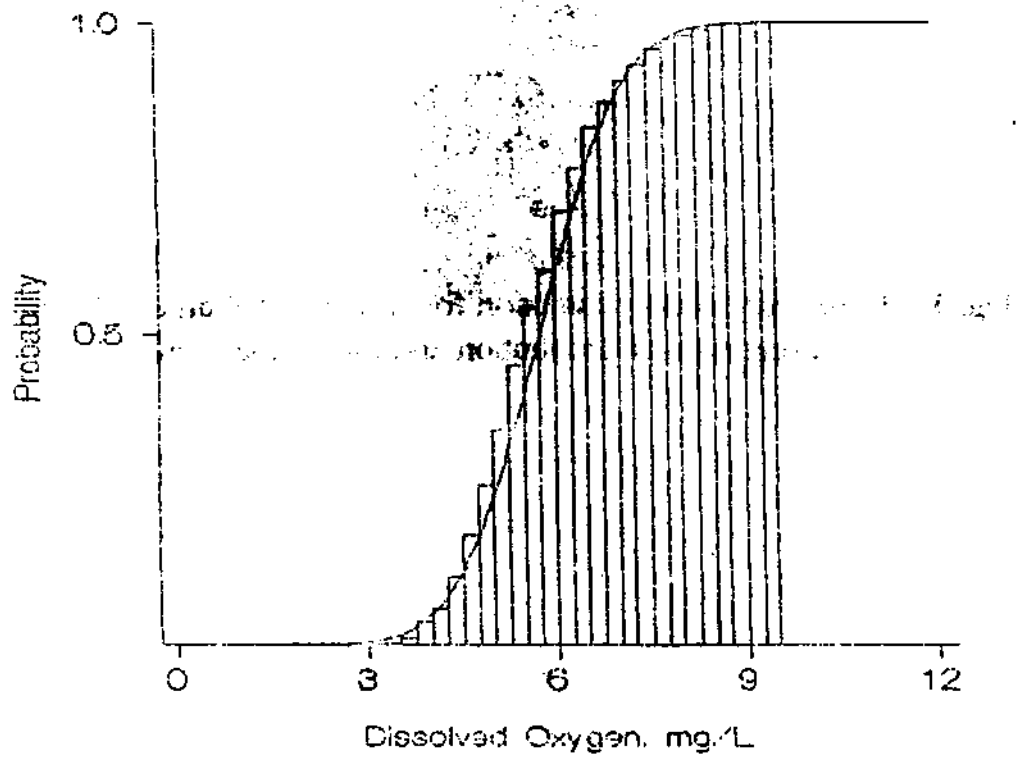

Fig.13 CPDF of DO Using First-Order Analysis and Log-Normal Distribution for DO

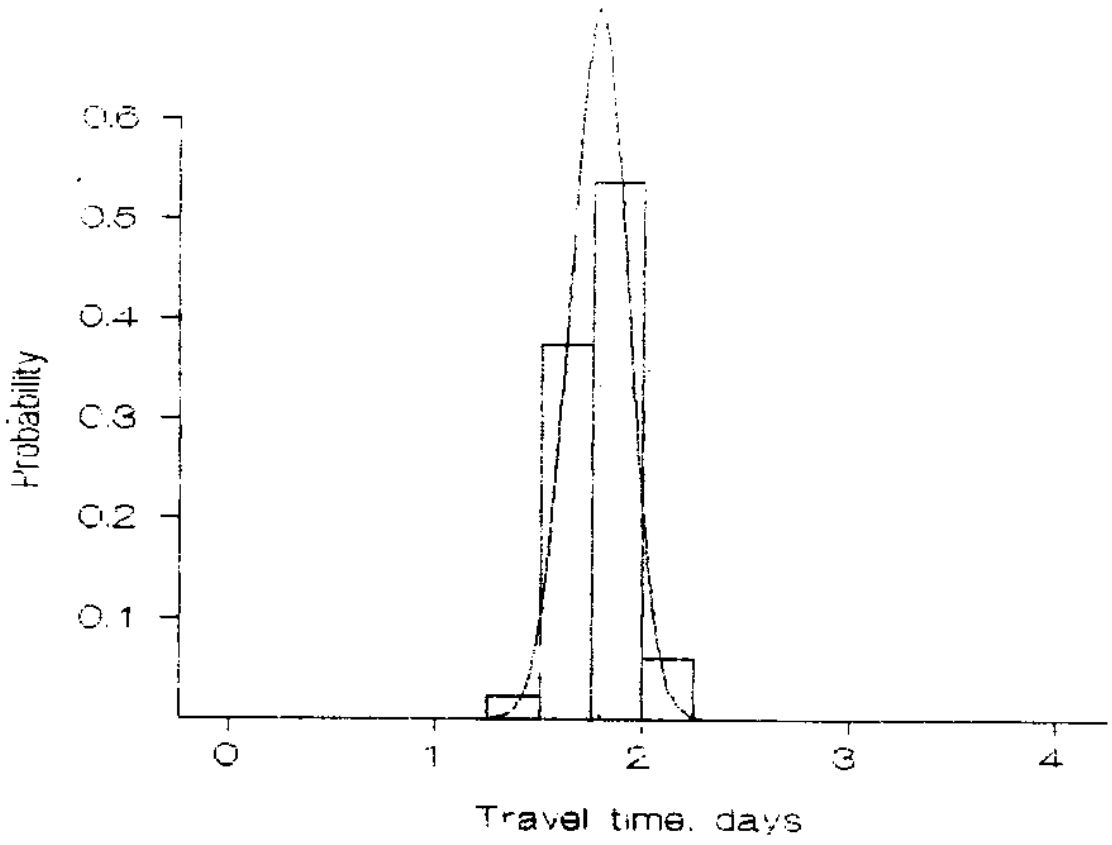

Fig. 14: Probability density Function for critical travel time obtained using normal distribution and First Order Analysis

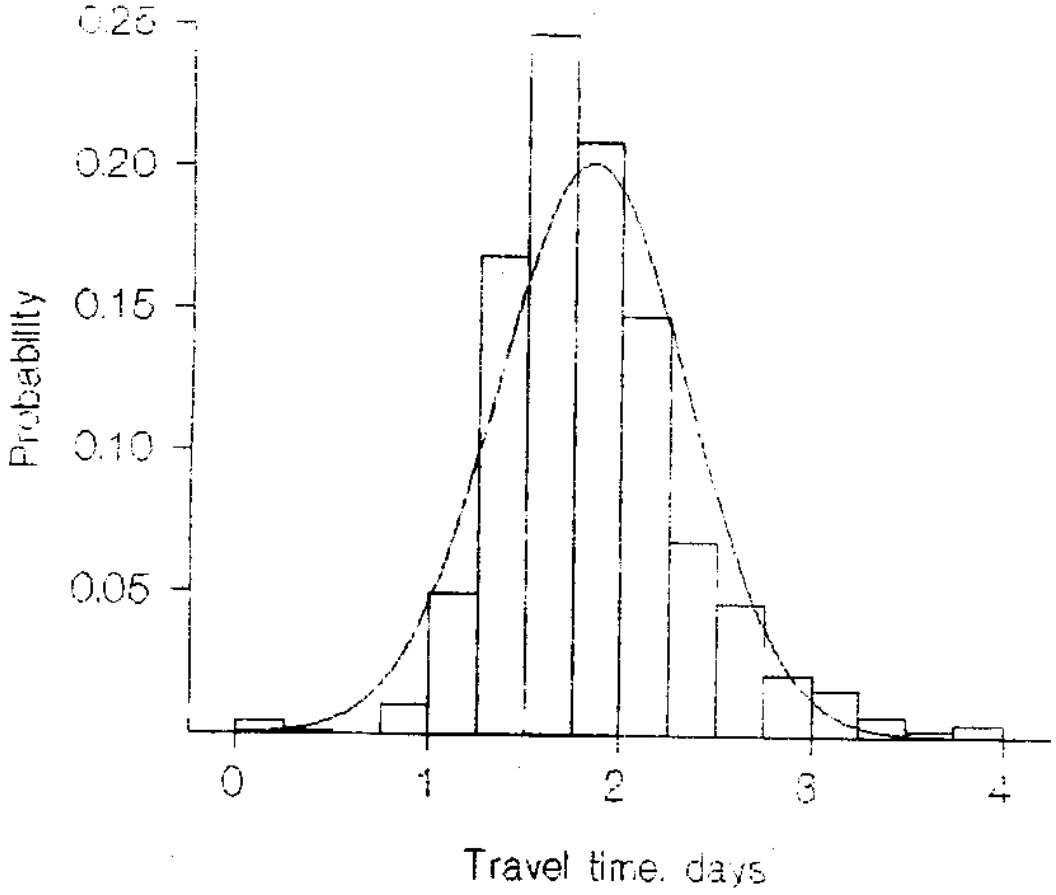

Fig.15: Probability density Function for critical travel time obtained using Monte Carlo Simulation (input to be Normally distributed)

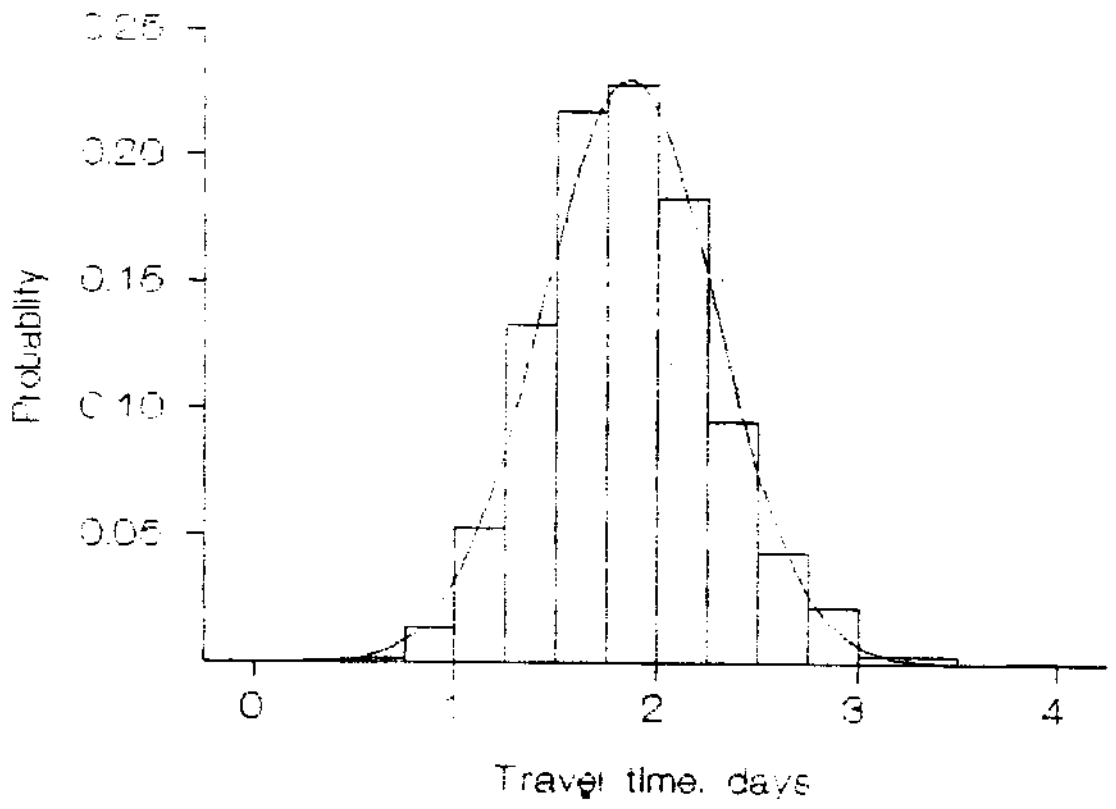

Fig.16: Probability density Function for critical travel time obtained using Monte Carlo Simulation (input to be Log-Normally distributed)

7.0 REFERENCES

Benjamin, J. R., and Cornell, C.A. (1970). "Probability, Statistics, and Decision for Civil Engineers", McGraw-Hill Book Co., Inc., New York, N.Y.

Burgess, S.J. (1979). "Analysis of Uncertainty in Flood plain Mapping, "Water Resources Bulletin, Vol. 15, No. 1, pp. 227-243.

Carter, R.W., and Anderson, I.E. (1963). "Accuracy of Current-Meter Measurements," American Society of Civil Engineers, Vol. 89, No. HY4, pp. 105-115

Churchill, M.A.,, Elmore, H.L., and Buckingham, R. A. (1962). "The Prediction of Stream Reaeration Rates," Journal of the Sanitary Engineering Division, ASCE, Vol. 88, No. SA4, pp. 1-46.

Chadderton, R. A., Miller, a.C., and McDonnell, A.J. (1982). "Uncertainty analysis of dissolved oxygen model." J. Envir. Engrs. Div., ASCE, 108(5), 1003-1012.

Chadderton, R.A., and Miller, A.C. (1980). "Analysis of Uncertainty in Floodplain Mapping," Water Resources Bulletin, Vol. 16, No. 4, pp. 752-754.

Fair, G.M., Geyer, J.C., and Okun, D.A. (1968). Water and Wastewater Engineering, Vol. 2, John Wiley and Sons, Inc., New York, N.Y., pp. 32-22 to 33-29.

Kothandaraman, V., and Ewing, B.B. (1969). "A probabilistic analysis of dissolved oxygen-biochemical oxygen demand relationship in stream." J. Water Poll. Cont. Fed., 41(2), R73-R90.

Loucks, D.P., and Lynn, W.R. (1966). "Probabilistic models for predicting stream quality," Water Resour. Res., 2(3), 593-605.

Padgett, W.J., and Rao, A.N.V. (1979). "Estimation of BOD and DO probability distribution." J. Envir. Engrg. Div., ASCE, 105(3), 525-533.

Streeter, H.W., and Phelps, E.B. (1925). "A study of the pollution and natural purification of the Ohio River," Public Health Bulletin, U.S. Public Health Service, Washington, D.C.

O'Connor and Dobbins (1958). "Mechanism of Reaeration in Natural Streams," Transactions, ASCE, Vol. 123, pp- 641-684.

Owens, M., Edwards, R. W., and Gibbs, J.W. (1964). "Some Reaeration Studies in Streams," International Journal Air and Water Pollution, Vol. 8, pp. 469-486.

Brown, L. C., and T.O. Barnwell, Jr.(1987). " The enhanced stream water quality models QUAL2E and QUAL2E-UNcas: Documentation and user mannual report". Environ. Res. Lab., U. S. Environ. Prot. Agency, Washington, D.C..

Burges, S.J. and Lettenmair, D.P.(1975). " Probabilistic methods in stream quality management, Water Resources Bulletin, 1, 11, 115-130.

Hornberger, G. M. (1980). "uncertainty in dissolved oxygen prediction due to variability in algal photosynthesis". Water Research, 14(4), 355-361.

Patel, J. K., Kapadia, C. H., and Owen, D. B.(1976). " Handbook of statistical distributions. John Wiley and Sons, Inc., New York, N.Y..

DIRECTOR	**:**	**Dr. S.M. Seth**
Div. Head	**:**	**Dr. K.K.S. Bhatia**
Study Group	**:**	**Aditya Tyagi, Dr. KKS Bhatia, and M. K. Sharma**