STOCHASTIC MODEL FOR POLLUTION AND DISSOLVED OXYGEN IN STREAMS

Cover designed by David J. E. Greene and John Wisnosky

STOCHASTIC MODEL FOR POLLUTION AND DISSOLVED OXYGEN IN STREAMS

A STOCHASTIC MODEL FOR POLLUTION
AND
DISSOLVED OXYGEN IN STREAMS

urnell
Richard P. Thayer
Graduate Student
Department of Statistics

Richard G. Krutchkoff
Associate Professor
Department of Statistics

Water Resources Research Center

Virginia Polytechnic Institute

Blacksburg, Virginia

August 1966

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the staff at Virginia Polytechnic Institute, in particular, to Professors, William Parsons, Department of Civil Engineering, and Boyd Harshbarger, Department of Statistics, for their suggestions and cooperation.

The research presented here was financed jointly by the Office of Resource Development, Bureau of State Services, (Environmental Health), and the Virginia Agricultural Experiment Station.

This publication was partially financed by the Department of the Interior under the terms of the Water Resources Research Act of 1964, Public Law 88-379.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 4
  1.1 The General Problem 4
  1.2 Previous Developments 7
  1.3 Scope and Aims 12
II. THEORETICAL RESULTS 14
  2.1 Stochastic Development 14
  2.2 Fixed Point Initial Conditions. 26
  2.3 Computer Program Usage 34
III. OTHER INITIAL CONDITIONS 37
  3.1 Variable Stream Conditions 37
  3.2 Measured Variability Initial Conditions 42
IV. APPLICATION TO SACRAMENTO RIVER. 50
  4.1 Description 50
  4.2 Analysis of the Data 54
  4.3 Applications to Future Situations . . . 65
V. LABORATORY TESTS 79
  5.1 Apparatus 79
  5.2 Estimation of Parameters. 81
  5.3 First Trial Run 87
  5.4 Second Trial Run. 98
VI. CONCLUSIONS 107
  6.1 Uses of the Results 107
  6.2 Conclusions 108
APPENDIX I COMPUTER PROGRAMS 110
APPENDIX II MATHEMATICS 118
APPENDIX III METHODS OF MEASUREMENT 124
REFERENCES 130

LIST OF FIGURES AND TABLES

		Page
Figure 1.2.1	Dissolved Oxygen versus Time of Travel .	11
Table 2.3.1	Arrangement of Data Card	36
Table 3.2.1	Arrangement of Data Card	49
Table 4.1.1	Sacramento River Data	53
Figure 4.2.1	Mean and Lower Confidence Curve for DO .	58
Table 4.2.1	Number of Points Below Confidence Bound.	59
Figure 4.2.2	Predicted and Observed Variances	61
Figure 4.2.3	Control Chart for Ranges	64
Table 4.3.1	Computer Printout for Future Conditions.	67
Figure 4.3.1	Pollution and Dissolved Oxygen Mass Densities for t = 2	76
Figure 4.3.2	Pollution and DO Means and Confidence Limits	77
Table 5.2.1	Data for Estimating the Parameter K_1+K_3	83
Table 5.2.2	Estimation of the Parameter K_1	84
Table 5.2.3	Data for Estimating the Parameter K_2 . .	86
Figure 5.3.1	Predicted and Observed DO Values - Run 1	89
Figure 5.3.2	Predicted and Observed COD Values - Run 1	90
Table 5.3.1	Sample Pollution Variances - Run 1 . . .	92
Figure 5.3.3	Predicted and Observed DO Variances - Run 1	95
Figure 5.3.4	Control Chart for Ranges - Run 1	96
Figure 5.3.5	Predicted and Observed DO Variances - Run 1	97
Figure 5.4.1	Predicted and Observed COD Values - Run 2	101
Figure 5.4.2	Predicted and Observed DO Values - Run 2	102
Table 5.4.1	Sample Pollution Variances - Run 2 . . .	103
Figure 5.4.3	Predicted and Observed DO Variances - Run 2	105
Figure 5.4.4	Control Chart for Ranges - Run 2	106

I. INTRODUCTION

1.1 The General Problem

The problem of the pollution of the rivers and estuaries of this nation and the world is now receiving considerable attention, and rightly so. Many rivers are so grossly polluted that there is scum on the surface and the odor of methane and hydrogen sulfide is noticeable. Consequently, there is a danger to health. For example, part of the Hudson River in New York is so foul that only eels can live in it. The lower Mississippi River is full of fish that have died. Through a great effort over a long period of time, the sewage load on the Potomac River has been reduced. The river has improved greatly and the more desirable species of fish are beginning to return.

With the population increasing at such a rapid rate, the problem becomes more severe. More sewage and industrial pollution are added to the rivers and more water is needed for domestic and industrial use. Unfortunately, the amount of water available cannot be increased and therefore it is a question of using wisely the water that is available. At present, the money being spent on sewage treatment plants is hardly enough to keep up with the increased population, which implies that nothing is being done to reduce the backlog. The Public Health Service estimates that through the next decade it will take 800 million dollars a year to catch up on the backlog - building of new plants, replacement for old plants and expansion of present facilities.

What are the functions that a river serves? Initially, water is taken from the river and given a degree of treatment depending on its intended use. The treatment may include removing settleable material, chlorinating to kill bacteria which might be pathogenic and removing undesirable tastes, odors, or color. Then the water is used either domestically or industrially and part of the water is returned to the river, with or without additional treatment. The quality of the water returned to the river is usually worse than when it was withdrawn.

Secondly, the rivers are the source of fish for the commercial fishing industry. With increasing degradation of the stream, the yield of fish decreases. For example, the economy of the Pacific Northwest is strongly dependent on the salmon industry. The salmon return from the ocean to the headwaters where they were spawned to lay their eggs. Any gross pollution along the route could deter the fish and there would be fewer salmon in the river.

Finally, the public depends on the rivers and lakes for recreation. Sport fishing and boating are popular pastimes. Swimming requires that the water be appealing and free from disease-causing organisms. A water course should be clean for its aesthetic beauty. For these and other reasons it is necessary to keep the rivers clean. Failure causes health hazards, requires expensive purification before water can be used, kills or drives away the better fish and spoils recreational opportunities.

On the other hand, rivers have a capacity for assimilating pollution and thus saving on the cost of waste treatment. For the sewage of a city it is not too difficult to remove the majority of the pollution by allowing it to settle out in a quiescent tank. To get a high degree of purity of the effluent requires considerably more equipment and considerably higher cost. When pollution is put into a river, the bacteria present in the river decompose the complex dead organic matter into simpler compounds. The complex carbon molecules are broken down first with the release of carbon dioxide. After a time the complex nitrogen molecules are broken down, yielding nitrates ultimately. These simple compounds are stable, non-offensive and harmless. The bacteria use the oxygen in the water in their feeding process.

The amount of dissolved oxygen in the water is critical because the maximum solubility of oxygen in a stream is relatively low; the concentration is only about 9 parts per million at 70°F. The solubility decreases with increasing temperatures, so that in the warmer months, when the solubility is least and the flows are lowest, the danger is greatest.

Complicating the problem also is the fact that industries use the water for cooling and return it to the stream at a higher temperature than they take it. This is called thermal pollution. The temperature of the river may be greater than would be expected under natural conditions; hence there is a lower maximum dissolved oxygen level.

There is always some pollution in a river. It may be due to the remainder from some large source upstream or it may be just small bits added along the water course. The result is that the dissolved oxygen is never able to maintain its maximum.

Now that the maximum has been discussed, what are the minimum desired levels of dissolved oxygen? Although certain fish such as salmon require more, it is generally true that fish need roughly 5 ppm of dissolved oxygen.

It is found that the range from the maximum present to the minimum desired is always small. This fact means that dissolved oxygen levels must be carefully observed because the concentration can go below desired levels very quickly.

1.2 Previous Developments

Before discussing the previous work, it is necessary to discuss explicitly what is meant by pollution and dissolved oxygen. The dissolved oxygen concentration, denoted by DO, in a river is measured in parts per million, ppm. These units of concentration are equivalent to the units milligrams per liter, abbreviated mg/l, and the two will be used interchangeably. The amount of pollution is indicated by the ultimate Biochemical Oxygen Demand, BOD, and is measured in units of ppm or mg/l. Essentially, the BOD is the amount of oxygen consumed by the bacteria in stabilizing the organic matter in a stream sample. Temperature, amounts of sample and other conditions are also specified for this determination. One way of thinking of these concepts is to consider the BOD as that which is owed and must be paid at a later time. The dissolved oxygen or DO is that which is on hand at the present time. One is not content just to have a high DO if the BOD is also high. Therefore, both quantities must be considered when evaluating the condition of a river.

It should be remarked at this time that any change in DO or BOD, such as an increase in pollution from an outside source, should be measured in terms of its effect on the BOD and DO of the main river.

There are many different forms of pollution and their effects are varied. The concern here is with organic pollution only and its effect on the dissolved oxygen. For example, industrial chemical pollution and its effect on the river are not considered. Some industrial chemicals even in very minute amounts have a harmful effect on fish. Also, undesirable colors from industrial wastes are not considered unless, of course, these happen to be organic dyes.

For three decades sanitary engineers have been studying the factors affecting the BOD and DO of the stream as it pursues its downstream course. In the earliest work Streeter and Phelps (11) considered only the following factors; the pollution decrease by the action of the bacteria, the oxygen decrease for the same reason and the oxygen increase by reaeration from the atmosphere. These still are the major factors, but others have been included in more recent work.

Very little work has been done on this problem from a statistical point of view. To this author's knowledge most of the attempts thus far have been of a multiple regression and correlation nature.

One of the most recent works of a non-statistical nature was published by Dobbins (3) in which he combines the work of many previous authors. Because many of the ideas and assumptions of Dobbins' (3) article are included in later chapters, some of his work will be briefly discussed here. Three assumptions must be made at the outset:

1. The stream flow is steady and uniform.
2. The process for the stretch as a whole is a steady-state process, the conditions at every cross-section being unchanged with time.
3. The BOD and DO are uniformly distributed over each cross-section, thus permitting the equations to be written in the usual one-dimensional form.

Assuming a constant velocity of the stream, the rate of change of the pollution, L, with respect to time of travel, t, is given, according to Dobbins (3), by the deterministic differential equation

$$\frac{dL}{dt} = L_a - (K_1 + K_3)L \qquad (1.2.1)$$

The terms are as follows:

1. L_a is the rate of addition of pollution along the route and is uniform (ppm/time). This term includes minor sources only; a large increase would require separate calculation from that point on.
2. K_1L is the rate of decrease in pollution due to the action of the bacteria, the rate of decrease being proportional to the amount of pollution present. Of course the units on K_1 are day^{-1}.
3. K_3L is the rate of decrease in pollution due to sedimentation of solids to the bottom and adsorption of the solids to the rocks and debris. In addition, this includes other losses of pollution which do not cause a corresponding loss in dissolved oxygen. The rate of decrease is also proportional to the amount of pollution present and the units of K_3 are per day.

For an initial condition of L_A ppm at time $t = 0$, the differential equation has the solution

$$L = L_A e^{-(K_1 + K_3)t} + \frac{L_a}{K_1 + K_3}(1 - e^{-(K_1+K_3)t}) \qquad (1.2.2)$$

The graph of pollution versus time of travel is a simple exponential decay from the maximum L_A to the ultimate value $L_a/(K_1 + K_3)$.

Similarly for the dissolved oxygen, the rate of change of the dissolved oxygen, C, with respect to time of travel is given according to Dobbins (3) by the deterministic differential equation

$$\frac{dC}{dt} = K_2(C_N' - C) - K_1L - D_B \qquad (1.2.3)$$

The parameters in the equation are as follows:

1. C_N' is the maximum oxygen solubility concentration under the given conditions of temperature and barometric pressure. Its units are ppm.
2. K_2C is the increase in dissolved oxygen due to the reaeration of the stream by the oxygen in the atmosphere. It is not a purely physical process because it can be affected adversely by solutes in the water. The rate is proportional to the dissolved oxygen deficit and it is in units of per day.
3. K_1L is the decrease in dissolved oxygen due to the action of the bacteria. The parameter K_1 is the same K_1 as discussed above, the rate being proportional to the amount of pollution present. The rate is therefore independent of the concentration of the dissolved oxygen. Strictly speaking this assumption is not valid below about 1 ppm DO, but since values that low are out of the range of interest, this is not a problem.
4. D_B is the decrease in DO due to the benthal demand, that is, the oxygen used by the bacteria on the bottom of the river. It also includes the effects of algae. The rate is independent of the DO present and it is expressed in units of ppm/day.

The pollution, L, in equation (1.2.3) is replaced by the value obtained in equation (1.2.2). For an initial condition of C_o at $t = 0$, the differential equation has the solution

$$C = C_N'(1 - e^{-K_2t}) - \frac{K_1\left(L_A - \frac{L_a}{K_1 + K_3}\right)}{K_2 - (K_1 + K_3)}\left(e^{-(K_1+K_3)t} - e^{-K_2t}\right)$$

$$-\left(\frac{D_B}{K_2} + \frac{K_1L_a}{K_2(K_1+K_3)}\right)(1 - e^{-K_2t}) + C_oe^{-K_2t} \qquad (1.2.4)$$

The above function is approximately of the form graphed below.

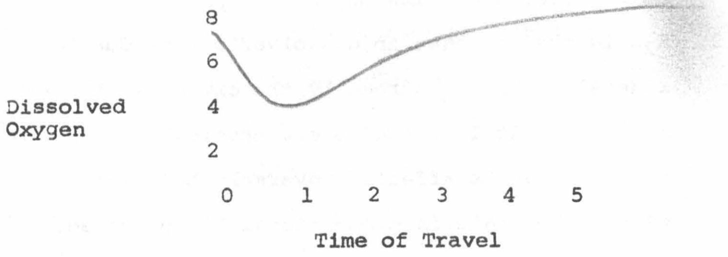

Figure 1.2.1 Dissolved Oxygen Versus Time of Travel

These two derived mean value functions are only applicable in the major regions of interest. For example, it is possible to make C negative by choosing a very large L_A. However, this equation is only applied in practice for values of the parameters that give reasonable values for C. Also, K_2 is almost always greater than $K_1 + K_3$, so that it is not worth considering $K_2 = K_1 + K_3$ in which case C would be indeterminate.

Some authors have included an additional factor known as longitudinal dispersion. This factor is a measure of the nonuniformity with which the stream moves, or in other words, is the amount of short-circuiting that takes place as the stream progresses. Dobbins (3) includes the term and then concludes that this factor has very little effect on the results and hence drops it. Longitudinal dispersion becomes somewhat more important when the equations are applied to estuaries.

Only the average conditions of a stream have been considered in the previous work and the possible variability about the average has not been taken into account; hence, the usefullness of the results is limited. The purpose of the present work is to find the probability distribution of the amounts of both pollution and dissolved oxygen in polluted streams as a function of time of travel. It is hoped that the mean values will be consistent with Dobbins (3) findings. The distributions will allow one to see what proportion of the time the pollution and dissolved oxygen are outside of desirable limits. Hopefully, these results will lead to more realistic restrictions on the use of streams for disposing of wastes than merely a knowledge of the mean permits.

1.3 Scope and Aims

The objective of this dissertation will be to consider the problem of pollution and dissolved oxygen in rivers from a stochastic point of view rather than from a deterministic point of view. Many of the assumptions used here are those accepted and applied in the field of sanitary engineering. One additional assumption will be needed. It is proposed to view the problem as a Stochastic Birth and Death Process with the BOD and DO being increased and decreased only by small amounts in a very short time interval. With this assumption a stochastic model will be developed.

The model will provide a joint density function for both pollution and dissolved oxygen when initial conditions are assumed. There are many different initial conditions possible when one considers variability. There is basic river variability and the variability with which the initial pollution is added. Several possible initial conditions that could be encountered in practice and the resulting river profiles will be the subject matter of later chapters.

Although it will be possible to compute conditional distributions these are of no interest. The question to which this investigation is addressed is, given the initial conditions of both BOD and DO what is the joint probability distribution and what are the two marginal probability distributions at a point downstream? For example, in no case are we given a BOD measurement downstream and asked to find the DO distribution at the point. For the same reason, there is no real interest in covariances and correlations of the two variables.

II. THEORETICAL RESULTS

2.1 Stochastic Development

As in the work of Dobbins (3) discussed previously, it will be assumed that the factors which can cause the concentration of pollution and dissolved oxygen to change are as follows:

1. Pollution added at the rate L_a from an outside source.
2. Pollution decreased at the rate K_3L from sedimentation and adsorption.
3. Pollution and oxygen decreased at the rate K_1L from bacterial action.
4. Oxygen decreased at the rate D_B from the benthal demand.
5. Oxygen increased at the rate $K_2(C_N' - C)$ from the physical process of reaeration.

These factors were discussed in the previous chapter when they appeared in the differential equation solution to the problem.

It is assumed that the pollution or dissolved oxygen or both can change by a small amount Δ due to one of the factors described above. For example, the pollution may increase by Δ may increase by a multiple of Δ or it may not change, depending on whether or not pollution enters from an outside source, L_a. In many ways this process is of a discrete nature. The pollution from the sides, L_a, does not enter continuously, but in discrete quantities. Changes in the river depth, velocity and turbulence can cause discrete jumps in sedimentation and reaeration.

The value of Δ depends on the properties of the stream. The way in which Δ enters into the results will be evident later and at that time methods of estimation will be discussed.

The term state is defined as follows. A change of size Δ in the concentration will constitute a change of one state. Thus there are Δ ppm/state. The pollution state M will correspond to a concentration of L_M and the relationship between the two is

$$\text{state } M = (L_M \text{ ppm}) \div (\Delta \text{ ppm/state})$$

or simply

$$M = L_M \div \Delta$$

In a similar manner the state of dissolved oxygen N corresponds to a concentration of C_N ppm and the relationship between the two is

$$\text{state } N = (C_N \text{ ppm}) \div (\Delta \text{ ppm/state})$$

or simply

$$N = C_N \div \Delta$$

Throughout this development h will denote a small increment in time. The order of magnitude of h will be such that h raised to any positive power greater than one will be negligible in comparison to h. Any term involving h to a power greater than one will be denoted by o(h).

The assumptions made thus far in this section include the five listed at the beginning of the section and the additional one concerning the discreteness of the process in units of size Δ. The probabilities for the possible changes in state in time h will now be discussed. It is assumed that a change of more than one state by a factor in time h has probability o(h). The probability of a change of one state is assumed proportional to h. The probabilities of one change in state and no change in state for the five factors are as follows:

1. For the first factor the expected amount of increase in pollution due to L_a in the small time interval h is equal to $L_a h$. Let Pr(L increases by an amount Δ due to L_a in time h) = ah, where a is a proportionality constant. Then, E(change due to L_a in time h) = $\Delta p + o(h) + O(1-p-o(h))$. Solving for p one obtains $p = \frac{L_a h}{\Delta} + o(h)$. Pr(no increase due to L_a in time h) = $1 - \frac{L_a h}{\Delta} + o(h)$.
2. Expected decrease in pollution due to sedimentation in time h is $K_3 L_M h$. Pr(L decreases by an amount Δ due to K_3 in time h) = $\frac{K_3 L_M h}{\Delta} + o(h)$. Pr(no decrease due to K_3 in time h)

   $$= 1 - \frac{K_3 L_M h}{\Delta} + o(h).$$

3. Expected decrease in pollution and oxygen due to bacterial action in time h is $K_1 L_M h$. Pr(L and C decrease by an amount Δ due to K_1 in time h)

   $= \frac{K_1 L_M h}{\Delta} + o(h)$. Pr (no decrease due to K_1 in time h) = $1 - \frac{K_1 L_M h}{\Delta} + o(h)$.

4. Expected decrease in oxygen due to benthal demand in time h is $D_B h$. Pr(C decreases by an amount Δ due to D_B in time h) = $\frac{D_B h}{\Delta} + o(h)$. Pr(no decrease due to D_B in time h) = $1 - \frac{D_B h}{\Delta} + o(h)$.
5. Expected increase in oxygen due to reaeration in time h is $K_2(C_N' - C_N)h$. Pr(C increases by an amount Δ due to K_2 in time h) = $\frac{K_2(C_N' - C_N)h}{\Delta} + o(h)$.

   Pr(no increase due to K_2 in time h)

   $$= 1 - \frac{K_2(C_N' - C_N)h}{\Delta} + o(h).$$

The assumptions made concerning the expected amounts of change are simply the rate assumptions of Dobbins (3) multiplied by the time interval h.

The joint probability of finding the system with the pollution in state M and the dissolved oxygen in state N at time t is denoted by $Pr_{M,N}(t)$. Then $Pr_{M,N}(t + h)$ can be expressed in terms of probabilities at time t in the ways listed below. [When one factor changes two states o(h) will be used because this probability has already been assumed to be of this order. For a probability where two or more factors change state o(h) will be used, because the product of the probabilities will have h to a power greater than one]

1. At time t the pollution is in state M - 1 and the dissolved oxygen is in state N. In the interval h the pollution state increases by one and the dissolved oxygen state remains unchanged. The probability of this event occurring is

$$Pr_{M-1,N}(t) \left[\left(\frac{L_a h}{\Delta} + o(h)\right)\left(1 - \frac{K_3 L_{M-1} h}{\Delta} + o(h)\right)\right.$$

$$\left(1 - \frac{K_1 L_{M-1} h}{\Delta} + o(h)\right)\left(1 - \frac{D_B h}{\Delta} + o(h)\right)$$

$$\left.\left(1 - \frac{K_2 (C_N' - C_N) h}{\Delta} + o(h)\right) + o(h)\right] =$$

$$Pr_{M-1,N}(t) \left[\frac{L_a h}{\Delta} + o(h)\right].$$

2. At time t the states are M + 1 and N. In time h the pollution decreases. The probability of this event occurring is

$$Pr_{M+1,N}(t) \left[\left(\frac{K_3 L_{M+1} h}{\Delta} + o(h)\right)\left(1 - \frac{L_a h}{\Delta} + o(h)\right)\right.$$

$$\left(1 - \frac{K_1 L_{M+1} h}{\Delta} + o(h)\right)\left(1 - \frac{D_B h}{\Delta} + o(h)\right)$$

$$\left.\left(1 - \frac{K_2 (C_N' - C_N) h}{\Delta} + o(h)\right) + o(h)\right]$$

$$= Pr_{M+1,N}(t) \quad \frac{K_3 L_{M+1} h}{\Delta} + o(h)].$$

3. At time t the states are M + 1 and N + 1. In time h the pollution and the dissolved oxygen decrease. The probability is

$$Pr_{M+1,N+1}(t) \quad [(\frac{K_1 L_{M+1} h}{\Delta} + o(h))(1 - \frac{L_a h}{\Delta} + o(h))$$

$$(1 - \frac{K_3 L_{M+1} h}{\Delta} + o(h))(1 - \frac{D_B h}{\Delta} + o(H))$$

$$(1 - \frac{K_2 (C_{N'} - C_{N+1}) h}{\Delta} + o(h)) + o(h)] =$$

$$Pr_{M+1,N+1}(t) \quad \frac{K_1 L_{M+1} h}{\Delta} + o(h)].$$

4. At time t the states are M and N + 1. In time h the dissolved oxygen decreases. The probability of this event occurring is

$$Pr_{M,N+1}(t) \quad [(\frac{D_B h}{\Delta} + o(h))(1 - \frac{L_a h}{\Delta} + o(h))$$

$$(1 - \frac{K_3 L_M h}{\Delta} + o(h))(1 - \frac{K_1 L_M h}{\Delta} + o(h))$$

$$(1 - \frac{K_2 (C_{N'} - C_{N+1}) h}{\Delta} + o(h)) + o(h)] =$$

$$Pr_{M,N+1}(t) \quad \frac{D_B h}{\Delta} + o(h)].$$

5. At time t the states are M and N - 1. In time h the dissolved oxygen increases. The probability of this event occurring is

$$Pr_{M,N-1}(t) \quad (\frac{K_2 (C_{N'} - C_{N-1}) h}{\Delta} + o(h))(1 - \frac{L_a h}{\Delta} + o(h))$$

$$(1 - \frac{K_3 L_M h}{\Delta} + o(h))(1 - \frac{K_1 L_M h}{\Delta} + o(h))$$

$$(1 - \frac{D_B h}{\Delta} + o(h)) + o(h)] =$$

$$Pr_{M,N-1}(t) \quad \frac{K_2 (C_{N'} - C_{N-1}) h}{\Delta} + o(h)].$$

6. At time t the pollution is in state M and the dissolved oxygen is in state N. In the interval h no change takes place. The probability of this event occurring is

$$\begin{aligned}
&Pr_{M,N}(t)\,[(1 - \frac{L_a h}{\Delta} + o(h))(1 - \frac{K_3 L_M h}{\Delta} + o(h)) \\
&(1 - \frac{K_1 L_M h}{\Delta} + o(h))(1 - \frac{D_B h}{\Delta} + o(h)) \\
&(1 - \frac{K_2(C_{N'} - C_N)h}{\Delta} + o(h)) + o(h)] \;= \\
&Pr_{M,N}(t)\;[1 - \frac{L_a h}{\Delta} - \frac{K_3 L_M h}{\Delta} - \frac{K_1 L_M h}{\Delta} - \frac{D_B h}{\Delta} \\
&- \frac{K_2(C_{N'} - C_N)h}{\Delta} + o(h)].
\end{aligned}$$

The coefficients of $Pr_{M-1,N+1}(t)$, $Pr_{M+1,N-1}(t)$ and $Pr_{M-1,N-1}(t)$ are o(h) and also all of the probabilities of the form $Pr_{M+a,N+b}(t)$ where either a or b is 2, 3, 4, ... have coefficients o(h). Throughout the above it has been assumed that the factors are independent so that the joint probabilities could be written as the product of marginal probabilities.

Combining these disjoint probabilities and ignoring o(h), one obtains the difference equation

$$\begin{aligned}
&Pr_{M,N}(t+h) = [\,1 - \frac{L_a h}{\Delta} - K_3 Mh - K_1 Mh - \frac{D_B h}{\Delta} - \\
&K_2(\frac{C_{N'}}{\Delta} - N)h\,]\; Pr_{M,N}(t) + \frac{L_a h}{\Delta} Pr_{M-1,N}(t) + \\
&K_3(M+1)hPr_{M+1,N}(t) + K_1(M+1)hPr_{M+1,N+1}(t) + \\
&\frac{D_B h}{\Delta} Pr_{M,N+1}(t) + K_2(\frac{C_{N'}}{\Delta} - (N-1))hPr_{M,N-1}(t). \qquad (2.1.1)
\end{aligned}$$

It is assumed that $Pr_{v,-1}(t)$ and $Pr_{-1,v}(t)$ are zero for all v. In this equation L_M/Δ was replaced by M and C_N/Δ was replaced by N. This is a difference equation in M, N and the continuous time variable t taken in units of size h.

Equation (2.1.1) gives the difference equation for the probability densities. It is desired now to find the form of $Pr_{M,N}(t)$ explicitly. Given $Pr_{M,N}(0)$, an initial probability distribution, one could find $Pr_{M,N}(t)$ where t is the product of an integer n times h. One could, proceed by progressing in units of size h, find the probability distribution at h, 2h, ..., nh. This process is very tedious especially since h must be small so that o(h) can be neglected. These problems can be avoided by the following method.

In equation (2.1.1) in which t is a continuous variable subtract $Pr_{M,N}(t)$ from both sides, divide both sides by h and take the limit as h approaches zero. The result is the differential-difference equation

$$\frac{d}{dt}Pr_{M,N}(t) = [- \frac{L_a}{\Delta} - K_3 M - K_1 M - \frac{D_B}{\Delta} - K_2(\frac{C_{N'}}{\Delta} - N)$$

$$Pr_{M,N}(t) + K_3(M + 1)Pr_{M+1,N}(t) + K_1(M + 1)Pr_{M+1,N+1}(t) +$$

$$\frac{L_a}{\Delta}Pr_{M-1,N}(t) + \frac{D_B}{\Delta}Pr_{M,N+1}(t) + K_2(\frac{C_{N'}}{\Delta} - (N-1))Pr_{M,N-1}(t).$$

(2.1.2)

The remainder of this section will deal with the probability generating function technique of finding a solution to equation (2.1.2). The probability generating function is defined as

$$P(S,R;t) = \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N}(t)S^M R^N.$$

Some of the basic generating function manipulations will be found in Appendix II.

Each differential-difference equation is multiplied by S^M and R^N, where M and N correspond to the subscripts of the probability on the left side of the equations. The set of equations is then summed over all values of both M and N yielding

$$
\begin{gathered}
\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}\frac{d}{dt}Pr_{M,N}(t)S^MR^N = \frac{L_a}{\Delta}\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}Pr_{M,N}(t)S^MR^N - \\
K_3\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}MPr_{M,N}(t)S^MR^N - K_1\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}MPr_{M,N}(t)S^MR^N - \\
\frac{D_B}{\Delta}\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}Pr_{M,N}(t)S^MR^N - K_2\frac{C_{N'}}{\Delta}\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}Pr_{M,N}(t)S^MR^N + \\
K_2\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}NPr_{M,N}(t)S^MR^N + \frac{L_a}{\Delta}\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}Pr_{M-1,N}(t)S^MR^N + \\
K_3\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}(M+1)Pr_{M+1,N}(t)S^MR^N + K_1\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}(M+1)\cdot \\
Pr_{M+1,N+1}(t)S^MR^N + \frac{D_B}{\Delta}\sum_{M=0}^{\infty}\sum_{N=0}^{\infty}Pr_{M,N+1}(t)S^MR^N + \\
K_2\frac{C_{N'}}{\Delta}\sum_{M=0}^{\infty}\sum_{N=1}^{\infty}Pr_{M,N-1}(t)S^MR^N - \\
K_2\sum_{M=0}^{\infty}\sum_{N=1}^{\infty}(N-1)Pr_{M,N-1}(t)S^MR^N. \qquad (2.1.3)
\end{gathered}
$$

The simplification of this equation involves algebraic manipulations which are discussed fully in Appendix II. The result is the partial differential equation

$$
\begin{gathered}
\frac{\partial P}{\partial t} = -\frac{L_a}{\Delta}P - K_3S\frac{\partial P}{\partial S} - K_1S\frac{\partial P}{\partial S} - \frac{D_B}{\Delta}P - K_2\frac{C_{N'}P}{\Delta} + \\
K_2R\frac{\partial P}{\partial R} + \frac{L_aSP}{\Delta} + K_3\frac{\partial P}{\partial S} + K_1\frac{1}{R}\frac{\partial P}{\partial S} + \frac{D_B}{\Delta}\frac{1}{R}P + K_2\frac{C_{N'}RP}{\Delta} - \\
K_2R^2\frac{\partial P}{R} - \frac{1}{R}\left(\frac{D_B}{\Delta} + K_1\frac{\partial}{\partial S}\right)P(S,0;t) \qquad (2.1.4)
\end{gathered}
$$

where $P = P(S,R;t)$.

Consider the term $\frac{1}{R}\left(\frac{D_B}{\Delta} + K_1\frac{\partial}{\partial S}\right)P(S,0;t)$. To see how this term arises it is necessary to review some assumptions, namely, that for the factors D_B and K_1 the transition probabilities are independent of the dissolved oxygen state. In particular, even if the dissolved oxygen is in state zero, there is a non-zero probability that the dissolved oxygen will decrease. This of course can not be true. In the term $P(S,0;t)$ which equals $\sum_{M=0}^{\infty} Pr_{M,0}(t)\ S^M$, one sums on the pollution variable M when the dissolved oxygen state is zero, that is, N=0. In (2.1.4) the term $\frac{1}{R}(\frac{D_B}{\Delta} + K_1 \frac{\partial}{\partial S})P(S,0;t)$ had to be added to complete the generating function and then subtracted so that the net change was zero.

The assumption of independence of D_B and K_1 of the dissolved oxygen state actually begins to lose its validity when the dissolved oxygen gets as low as 1 - 2 ppm. There is some dependence on the dissolved oxygen state at low levels. Sanitary engineers use this assumption realizing its limitations. No one has recommended a better assumption for low levels because the results are only applied when there is a reasonable amount of oxygen in the river. Therefore, for all practical purposes $Pr_{M,0}(t)$ is essentially zero for all M and t. The whole term $\frac{1}{R}(\frac{D_B}{\Delta} + K_1 \frac{\partial}{\partial S})P(S,0;t)$ is negligible and will be omitted from the partial differential equation. The justification of this assertion will be demonstrated in later chapters for particular cases.

Rewriting (2.1.4) without the negligible term gives

$$\frac{\partial P}{\partial t} = -\frac{L_a}{\Delta}P - K_3 S\frac{\partial P}{\partial S} - K_1 S\frac{\partial P}{\partial S} - \frac{D_B P}{\Delta} - K_2\frac{C_N'}{\Delta}P + K_2 R\frac{\partial P}{\partial R} +$$

$$\frac{L_a}{\Delta}SP + K_3\frac{\partial P}{\partial S} + K_1\cdot\frac{1}{R}\cdot\frac{\partial P}{\partial S} + \frac{D_B}{\Delta}\cdot\frac{1}{R}P + \frac{C_N'}{\Delta}RP - K_2 R^2\cdot\frac{\partial P}{\partial R}. \qquad (2.1.4a)$$

The equation can be further simplified by dividing through by P, letting z = ln P, and regrouping the terms. Equation (2.1.4a) then becomes

$$\frac{\partial z}{\partial t} - K_2 R(1-R)\frac{\partial z}{\partial R} - [K_3(1-S) + K_1(\frac{1}{R} - S)]\frac{\partial z}{\partial S} =$$

$$\frac{L_a}{\Delta}(1-S) - \frac{D_B}{\Delta}(\frac{1-R}{R}) + K_2\frac{C_N'}{\Delta}(1-R). \qquad (2.1.5)$$

The general theory of the solution of this type of partial differential equation is discussed in Appendix II. In particular, equation (2.1.5) is of the form of equation (II.1) in the Appendix. The correspondence is as follows:

$$y = z \qquad C = \frac{L_a}{\Delta}(1-S) - \frac{D_B}{\Delta}\left(\frac{1-R}{R}\right) + K_2\frac{C_{N'}}{\Delta}(1-R)$$

$$t = x_1 \qquad A_1 = 1$$

$$R = x_2 \qquad A_2 = -K_2R(1-R)$$

$$S = x_3 \qquad A_3 = -[K_3(1-S) + K_1(\tfrac{1}{R} - S)].$$

The subsidiary equations, (II.3) in Appendix II, are

$$\frac{dt}{1} = \frac{-\,dR}{K_2R(1-R)} = \frac{-dS}{K_3(1-S) + K_1(\frac{1}{R} - S)} =$$

$$\frac{dz}{\frac{L_a}{\Delta}(1-S) + \frac{D_B}{\Delta}\left(\frac{1-R}{R}\right) - K_2\frac{C_{N'}}{\Delta}(1-R)} \; . \qquad (2.1.6)$$

From these equations any three independent pairs are chosen for solution.

The first pair chosen will be the first and second because each one contains only the variable in the differential. After integration of the differential equation and solving for the constant of integration C, one obtains

$$C = \frac{R}{1-R}\, e^{K_2t} \; . \qquad (2.1.7)$$

The second pair chosen was the first and third terms in (2.1.6). Eliminate R from the third using (2.1.7) and then the only variables are t and S. Integration of the differential equation and subsequent elimination of C yields

$$B = \frac{-K_1}{K_2 - (K_1+K_3)}\left(\frac{1-R}{R}\right)e^{-(K_1+K_3)t} + (1-S)e^{-(K_1+K_3)t} \qquad (2.1.8)$$

where B is the constant of integration.

The third independent pair chosen was the second and fourth terms of (2.1.6). It is necessary to use (2.1.7) and (2.1.8.) to eliminate S and then t. On integrating the result and eliminating B. and C, one obtains

$$A = z + \frac{K_1 L_a}{\Delta[K_2 - (K_1+K_3)]K_2} \frac{1}{R} + \frac{L_a}{\Delta(K_1+K_3)} (1-S) -$$

$$\frac{K_1 L_a}{\Delta(K_1+K_3) \quad [K_1 - (K_1+K_3)]} \frac{(1-R)}{R} - \frac{D_B}{\Delta K_2} \frac{1}{R} - \frac{C_{N'}}{\Delta} \ln R \qquad (2.1.9)$$

where A is the new constant of integration.

The three equations (2.1.7), (2.1.8) and (2.1.9) correspond to the three equations (II.4) in Appendix II. The arbitrary function relating these solutions of the subsidiary equations is $A = \psi(B,C)$. In order to determine the ψ function explicitly, it is necessary to specify initial conditions. At time $t = 0$, it will be assumed that the probability generating function is known, i. e. $P(S,R;0)$ is specified. Specifying the form of $P(S,R;0)$ corresponds to equation (II.6) in the Appendix. Replacing z by $\ln P(S,R;t)$ in (2.1.9) and putting the values of A, B and C evaluated at $t = 0$ in the ψ function one obtains

$$\psi\left[(1-S) - \frac{K_1}{K_2 - (K_1+K_3)} \left(\frac{1-R}{R}\right), \frac{R}{1-R}\right] = \ln P)S,R;0) +$$

$$\frac{K_1 L_a}{\Delta K_2[K_2 - (K_1+K_3)]} \cdot \frac{1}{R} - \frac{K_1 L_a}{\Delta(K_1+K_3) \quad [K_2 - (K_1+K_3)]} \left(\frac{1-R}{R}\right) +$$

$$\frac{L_a}{\Delta(K_1+K_3)} (1-S) - \frac{D_B}{\Delta K_2} \cdot \frac{1}{R} - \frac{C_{N'}}{\Delta} \ln R. \qquad (2.1.10)$$

As in the Appendix, let

$$v = (1-S) - \frac{K_1}{K_2 - (K_1+K_3)} \left(\frac{1-R}{R}\right) \quad \text{and} \quad w = \frac{R}{1-R}$$

and then solve for R and S

$$S = (1-v) - \frac{K_1}{K_2 - (K_1+K_3)} \cdot \frac{1}{w} \text{ and } R = \frac{w}{1+w} .$$

Using these relationships, one can eliminate R and S from (2.1.10) to obtain

$$\psi(v,w) = \ln P(1 - v - \frac{K_1}{K_2 - (K_1+K_3)} \cdot \frac{1}{w}, \frac{w}{1+w} ; 0) +$$

$$\frac{K_1 L_a}{\Delta K_2 [K_2 - (K_1+K_3)]}(\frac{1+w}{w}) + \frac{L_a}{\Delta(K_1+K_3)} v -$$

$$\frac{D_B}{\Delta K_2}(\frac{1+w}{w}) - \frac{C_{N'}}{\Delta} \ln(\frac{w}{1+w}). \qquad (2.1.11)$$

This equation (2.1.11) is the form of the ψ function for any initial probability generating function. It corresponds to equation (II.8) in the Appendix. The particular case of interest is A = (B,C), therefore let v = B and w = c

$$A = [\ln P \quad 1 - B - \frac{K_1}{K_2 - (K_1+K_3)} \cdot \frac{1}{C}, \frac{C}{1+C}; 0] +$$

$$\frac{K_1 L_a}{\Delta K_2 [K_2 - (K_1 + K_3)]}(\frac{1+C}{C}) + \frac{L_a}{\Delta(K_1+K_3)} B - \frac{D_B}{\Delta K_2}(\frac{1+C}{C}) -$$

$$\frac{C_{N'}}{\Delta} \ln(\frac{C}{1+C}). \qquad (2.1.12)$$

The procedure to complete the development is to replace A, B and C by their respective equals in (2.1.7), (2.1.8) and (2.1.9). The equation can be simplified and solved for P(S,R;t) as a function of the general initial probability generating function. Given a specific initial generating function, P(S,R;t) is then completely known. However, this procedure becomes very tedious because the S and R in the initial generating function will themselves be functions of S and R when the substitutions are made.

A less involved procedure and the one that will be followed in subsequent chapters is to rearrange these last few steps.

1. As indicated in equation (2.1.12) let $S = 1 - B - \frac{K_1}{K_2 - (K_1+K_3)} \cdot \frac{1}{C}$ and $R = \frac{C}{1+C}$ in the given initial generating function and substitute the result in (2.1.12).
2. The new equation is a function of A, B and C, and these are replaced by their equals in (2.1.7), (2.1.8) and (2.1.9) respectively.
3. The result is simplified and solved for $P(S,R;t)$.

2.2 Fixed Point Initial Conditions

There are many different sets of initial conditions that one could assume, each of which corresponds to a different practical situation. The first condition to consider is: at time of travel equal to zero, both pollution and dissolved oxygen are at given fixed levels. These results will be used as a basis for other initial conditions.

For these initial conditions the dissolved oxygen is in state C_o/Δ with probability one and the pollution is in state L_A/Δ with probability one. The probability generating function is therefore

$$P(S,R;0) = S^{\frac{L_A}{\Delta}} R^{\frac{C_o}{\Delta}} \qquad (2.2.1)$$

Making the transformation of step 1 at the end of the previous section one obtains

$$P(1 - B - \frac{K_1}{K_2 - (K_1+K_3)} \cdot \frac{1}{C}, \frac{C}{1+C}; 0) =$$

$$(1 - B - \frac{K_1}{K_2 - (K_1+K_3)} \cdot \frac{1}{C})^{\frac{L_A}{\Delta}} (\frac{C}{1+C})^{\frac{C_o}{\Delta}} .$$

Substituting this result in equation (2.1.12), one obtains

$$A = \frac{L_A}{\Delta}\ln(1 - B - \frac{K_1}{K_2 - (K_1+K_3)}.\frac{1}{C}) + \frac{C_o}{\Delta}\ln(\frac{C}{1+C}) +$$

$$\frac{K_1 L_a}{\Delta K_2\ [K_2 - (K_1+K_3)]}(\frac{1+C}{C}) + \frac{L_a}{\Delta(K_1+K_3)}B - \frac{D_B}{\Delta K_2}(\frac{1+C}{C}) -$$

$$\frac{C_{N'}}{\Delta}\ln(\frac{C}{1+C}). \qquad (2.2.2)$$

In place of A, B and C one can substitute their equals as stated in step 2 of the previous section and simplify as stated in step 3 to obtain

$$P(S,R;t) = \exp[\frac{L_a K_1}{\Delta(K_1+K_3)(K_2 - (K_1+K_3)]}(1 - e^{-(K_1+K_3)t}).$$

$$(\frac{1-R}{R}) - \frac{L_a K_1}{\Delta K_2 [K_2 - (K_1+K_3)]}(1 - e^{-K_2 t})(\frac{1-R}{R}) -$$

$$\frac{L_a}{\Delta(K_1+K_3)}(1 - e^{-(K_1+K_3)t})(1-S) - \frac{D_B}{\Delta K_2}(1 - e^{-K_2 t})(\frac{1-R}{R}) \ .$$

$$[1 + (\frac{1-R}{R})e^{-K_2 t}]^{\frac{C_{N'}-C_o}{\Delta}} .R.^{\frac{C_{N'}}{\Delta}} [1 + \frac{K_1}{K_2 - (K_1+K_3)}.$$

$$(e^{-(K_1+K_3)t} - e^{-K_2 t})(\frac{1-R}{R}) - e^{-(K_1+K_3)t}(1-S)]^{\frac{L_A}{\Delta}}. \qquad (2.2.3)$$

This equation is the joint probability generating function for both pollution and oxygen. As stated in Appendix II this function could be expanded in a power series, the coefficient of $S^M R^N$ being $Pr_{M,N}(t)$, the probability that the pollution is in state M and the dissolved oxygen is in state N.

Consider next the marginal distribution for pollution, which is obtained by setting R equal to one in (2.2.3)

$$P(S,1;t) = \exp[-\frac{L_a}{(K_1+K_3)}(1 - e^{-(K_1+K_3)t})(1 - S) \ .$$

$$[1 - (1 - S)e^{-(K_1+K_3)t}]\frac{L_a}{\Delta} . \qquad (2.2.4)$$

This function can be expanded in a power series in S, the coefficient of S^M being $P_M(t)$, the probability that the pollution is in state M at time t. Thus,

$$P_M(t) = \exp [- \frac{L_a}{\Delta(K_1+K_3)} (1 - e^{-(K_1+K_3)t})] .$$

$$\sum_{i=0}^{M^*} \binom{\frac{L_A}{\Delta}}{i} (1 - e^{-(K_1+K_3)t})^{\frac{L_A}{\Delta} - i} (e^{-(K_1+K_3)t})^{i} .$$

$$\frac{[\frac{L_a}{\Delta(K_1+K_3)} (1 - e^{-(K_1+K_3)t})]^{M-i}}{(M-i)\,!} \qquad (2.2.5)$$

$$\text{where } M^* = \begin{cases} M & \text{for } M \leq L_A/\Delta \\ L_A/\Delta & \text{for } M > L_A/\Delta \end{cases}$$

A general discussion of moment generating functions can be found in Appendix II. The moment generating function for pollution M(0,t) is obtained by replacing S in (2.2.4) by e^{θ} giving

$$M(\theta, t) = \exp [- \frac{L_a}{\Delta(K_1+K_3)} (1 - e^{-(K_1+K_3)t})(1 - e^{\theta})] .$$

$$[1 - (1 - e^{\theta})e^{-(K_1+K_3)t}] \frac{L_A}{\Delta} . \qquad (2.2.6)$$

The first derivative of (2.2.6) evaluated at $\theta = 0$ gives the mean state of pollution as a function of time

$$\frac{L_a}{\Delta(K_1+K_3)} (1 - e^{-(K_1+K_3)t}) + \frac{L_A}{\Delta} e^{-(K_1+K_3)t} . \qquad (2.2.7)$$

Multiplying the mean state (2.2.7) by Δ gives the mean amount of pollution.

$$L(t) = \frac{L_a}{K_1+K_3} (1 - e^{-(K_1+K_3)t}) + L_A e^{-(K_1+K_3)t} . \qquad (2.2.8)$$

The significance of this result will be discussed later in this section.

The variance is obtained by taking the second derivative of the moment generating function, evaluating it at $\theta = 0$ and subtracting the square of the mean. After multiplying by Δ^2 the variance of the amount of pollution present is

$$\Delta\left[\frac{L_a}{K_1+K_3} + L_A e^{-(K_1+K_3)t}\right]\left[1 - e^{-(K_1+K_3)t}\right]. \qquad (2.2.9)$$

The form of the variance function is such that at $t = 0$ the variance is zero. As t increases the variance increases to a maximum and then decreases slowly to its asymptotic value $\frac{\Delta L_a}{K_1+K_3}$.

At $t = \infty$ the distribution becomes Poisson with the mean and variance of the states being equal. It is independent of the initial amount of pollution.

The marginal distribution for dissolved oxygen is found in a manner similar to that used for pollution. If S is set equal to unity in (2.2.3), one obtains

$$P(1,R;t) = \exp\left[\frac{L_a K_1}{\Delta(K_1+K_3)\left[K_2-(K_1+K_3)\right]}\left(\frac{1-R}{R}\right)\left(1-e^{-(K_1+K_3)t}\right)\right.$$

$$- \frac{L_{L_a}^{K}K_1}{\Delta K_2\left[K_2-(K_1+K_3)\right]}\left(\frac{1-R}{R}\right)\left(1-e^{-K_2 t}\right) + \left.\frac{D_B}{\Delta K_2}\left(\frac{1-R}{R}\right)\left(1-e^{-K_2 t}\right)\right].$$

$$\left[1 + \left(\frac{1-R}{R}\right)e^{-K_2 t}\right]^{\frac{C_{N'}-C_0}{\Delta}} \cdot R^{\frac{C_{N'}}{\Delta}} \cdot \left[1+ \frac{K_1}{K_2-(K_1+K_3)}\cdot\right.$$

$$\left.\left(e^{-(K_1+K_3)t} - e^{-K_2 t}\right)\left(\frac{1-R}{R}\right)\right]^{\frac{L_A}{\Delta}}. \qquad (2.2.10)$$

The power series expansion of this generating function is

$$P_N(t) = \exp\left[\frac{-L_a K_1}{\Delta(K_1+K_3)\left[K_2-(K_1+K_3)\right]}\left(1-e^{-(K_1+K_3)t}\right) + \right.$$

$$\frac{L_aK_1}{\Delta K_2\,[K_2-(K_1+K_3)]}\,(1-e^{-K_2t})-\frac{D_B}{\Delta K_2}(1-e^{-K_2t})\;].$$

$$\sum_{i=0}^{N^*}\;\sum_{j=0}^{N^{**}}\binom{\frac{L_A}{\Delta}}{i}\Bigl(1-\frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}-e^{-K_2t})\Bigr)^{\frac{L_A}{\Delta}-i}$$

$$\Bigl(\frac{K_1}{K_2-(K_j+K_3)}(e^{-(K_1+K_3)t}-e^{-K_2t})\Bigr)^{i}\binom{\frac{C_{N'}-C_0}{\Delta}}{j}(1-e^{-K_2t})^{\frac{C_{N'}-C_0}{\Delta}-j}$$

$$(e^{-K_2t})^{j}\;\Bigl[\frac{L_aK_1}{\Delta(K_1+K_3)\;\;K_2-(K_1+K_3)\,]}(1-e^{-(K_1+K_3)t})$$

$$+\;\frac{-L_aK_1}{\Delta K_2\,[K_2-(K_1+K_3)\,]}\;(1-e^{-K_2t})+\frac{D_B}{\Delta K_2}\,(1-e^{-K_2t})\Bigr]^{\frac{C_{N'}}{\Delta}-N-j-i}$$

$$\frac{1}{(\frac{C_{N'}}{\Delta}-N-j-i)!} \qquad (2.2.11)$$

where $N^* = \begin{cases} \frac{C_{N'}}{\Delta} - N & \text{for } \frac{C_{N'}}{\Delta} - N \le \frac{L_A}{\Delta} \\ \frac{L_A}{\Delta} & \text{for } \frac{C_{N'}}{\Delta} - N > \frac{L_A}{\Delta} \end{cases}$

and $N^{**} = \begin{cases} \frac{C_{N'}}{\Delta} - N - i & \text{for } \frac{C_{N'}}{\Delta} - N - i \le \frac{C_{N'}-C_0}{\Delta} \\ \frac{C_{N'}-C_0}{\Delta} & \text{for } \frac{C_{N'}}{\Delta} - N - i > \frac{C_{N'}-C_0}{\Delta} \end{cases}$.

This power series expansion in powers of R has exponents of R ranging from N' down to $-\infty$. The coefficient of R^N, namely $P_N(t)$, is the probability of being in state N at time t. The negative exponents are due to the fact that a term was dropped from the solution of the partial differential equation as discussed in Section 2.1. However, the coefficients when N is negative are negligibly small in cases where these results will be applied, so only the exponents from $\frac{C_{N'}}{\Delta}$ down to zero need be considered. This will be verified for particular cases in later chapters.

The moment generating function $M(\Phi,t)$ as explained in Appendix II is obtained by replacing R in (2.2.3) by e^{Φ}

$$M(\phi,t) = \exp(\{[\frac{L_aK_1}{\Delta(K_1+K_3)[K_2-(K_1+K_3)]}(1-e^{-(K_1+K_3)t})$$

$$-\frac{L_aK_1}{\Delta K_2[K_2-(K_1+K_3)]}(1-e^{-K_2t}) + \frac{D_B}{\Delta K_2}(1-e^{-K_2t}) \quad e^{-\phi}-1]\}).$$

$$[1+(e^{-\phi}-1)e^{-K_2t}]\frac{C_{N'}-C_0}{\Delta} \cdot e^{\phi}\frac{C_{N'}}{\Delta} \cdot [1+\frac{K_1}{K_2-(K_1+K_3)}$$

$$(e^{-K_1+K_3)t}-e^{-K_2t})(e^{-\phi}-1)]\frac{L_A}{\Delta}. \qquad (2.2.12)$$

The mean state of dissolved oxygen as a function of time is found by differentiating the moment generating function (2.2.12) and evaluating the result at $\phi=0$.

$$\frac{C_{N'}}{\Delta} - \frac{L_aK_1}{\Delta(K_1+K_3)[K_2-(K_1+K_3)]}(1-e^{-(K_1+K_3)t})$$

$$+\frac{L_aK_1}{\Delta K_2[K_2-(K_1+K_3)]}(1-e^{-K_2t}) - \frac{D_B}{\Delta K_2}(1-e^{-K_2t})$$

$$-(\frac{C_{N'}-C_0}{\Delta})e^{-K_2t} - \frac{L_A}{\Delta}\frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}-e^{-K_2t})$$

(2.2.13)

Multiplying both sides of (2.2.13) by Δ gives the mean amount of dissolved oxygen

$$C(t) = C_{N'} - \frac{L_aK_1}{(K_1+K_3)[K_2-(K_1+K_3)]}(1-e^{-(K_1+K_3)t})$$

$$+\frac{L_aK_1}{K_2[K_2-(K_1+K_3)]}(1-e^{-K_2t}) - \frac{D_B}{K_2}(1-e^{-K_2t})-(C_{N'}-C_0)e^{-K_2t}$$

$$-L_A\frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}-e^{-K_2t}). \qquad (2.2.14)$$

The significance of this result will be discussed at the end of this section.

The variance of the amount of dissolved oxygen present is

$$\Delta\{\frac{-L_aK_1}{K_2\,[K_2-(K_1+K_3)]}\,(1-e^{-K_2t}) + \frac{L_aK_1}{(K_1+K_3)\,[K_2-(K_1+K_3)]}$$

$$(1-e^{-(K_1+K_3)t}) + \frac{D_B}{K_2}(1-e^{-K_2t})+(C_N' - C_0)(1-e^{-K_2t})e^{-K_2t}$$

$$+ L_A\,\frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}-e^{-K_2t})(1- \frac{K_1}{K_2-(K_1+K_3)})$$

$$\cdot\,(e^{-(K_1+K_3)t}-e^{-K_2t}))\}\,. \qquad (2.2.15)$$

The variance is zero at t=0 as required. The variance at $t=\infty$ reduces to

$$\frac{\Delta L_aK_1}{K_2(K_1+K_3)} + \frac{D_B\,\Delta}{K_2}\,.$$

In between the variance rises quickly and "tapers off" slowly to its ultimate value.

The probability function, mean and variance for the dissolved oxygen are all functions of the initial state or amount of pollution. However, this initial effect becomes less important for large t.

Consider the mean pollution equation (2.2.8) and the mean dissolved oxygen equation (2.2.14). These two equations are exactly the same as the pollution and DO equations (1.2.2) and (1.2.4) respectively, obtained by Dobbins (3) using the deterministic differential equation approach. These results have been verified with actual data by sanitary engineers. The stochastic model obtained in this section has been derived from a completely different point of view, yet the results are compatible with other known and accepted results.

Since no previous theoretical work has given predicted variances, the only way in which the predicted variances of the present investigation can be checked is by obtaining actual river data and laboratory data. In later chapters data will be obtained and comparisons made between the observed and predicted variances.

Consider now the way in which Δ has entered into the results just obtained. The mean amount of both pollution and dissolved oxygen are independent of the choice of Δ . The variance of the amount of both pollution and dissolved oxygen are directly proportional to Δ. Thus the variance increases as Δ increases.

This quantity then should be estimated from the measurements made of the variance. When a stream survey is made, various types of measurements are taken which enable the investigators to estimate the stream parameters. During this survey repeated observations of both BOD and DO should be made at various points along the river. The time for these measurements should reasonably represent the period to which inferences are to be made. For example, if one can assume steady state conditions and if the results are to be applicable to summer conditions, then the measurements should be spread over this period. For each point there are predicted BOD and DO variances which, after the other parameters are estimated are functions of Δ . Hence, at each point at which these replicated observations are made there are two estimates of Δ, one from the BOD and the other from the DO. These Δ values should be reasonably close together, that is, within the limits of experimental error. From these estimates the value of Δ can be obtained by finding a weighted average. It turns out that the measurement error in the pollution is often large enough so that it masks the true variability. In this case the estimate of Δ can be obtained from the dissolved oxygen estimates of Δ . It is much more important to be able to predict dissolved oxygen variances than pollution variances, in any case.

2.3 Computer Program Usage

In order that these results be made more readily useful, a computer program was written to perform the arithmetic. Given the parameters the program loops on time and prints out all of the results discussed below in one day intervals from one to five days. The program can be easily adjusted to print out the result for other times of travel. Initial conditions or conditions at time $t = 0$ must be known in order to use the program. First the program computes the probability distribution for pollution and prints out the non-zero probabilities with their corresponding concentrations. Secondly for a given α confidence level the program sums the individual probabilities beginning with those corresponding to the upper concentrations of pollution. It stops and prints the concentration and cumulative probability for the nearest cumulative probability greater than α and the nearest cumulative probability less than α. This gives an upper confidence limit for pollution. Thirdly it evaluates the mean and variance functions for the amount of pollution and prints the results. This completes the pollution calculations.

Continuing with the same time of travel used for the pollution computations it begins the dissolved oxygen calculations. First it calculates the probability distribution and prints out the non-zero probabilities with their corresponding dissolved oxygen concentrations. Secondly it sums the probability mass densities beginning with the smallest dissolved oxygen concentration until it reaches the confidence limit α. Then it prints out the cumulative probability and the corresponding concentration for the nearest cumulative probability greater than α and the nearest cumulative probability less than α. This gives a lower confidence bound on the dissolved oxygen. Finally, it evaluates the mean and variance functions for the amount of dissolved oxygen and prints the results.

It was decided that an upper confidence limit is more desirable for pollution, for one wants to guard against having too much pollution at a certain point. Also it is obvious that a lower confidence bound is needed for dissolved oxygen because concentrations that are too low are detrimental to the fish.

The main program, written in Fortran IV, is listed as Program I in Appendix I. The Exponential Subroutine, the Binomial Subroutine and the Series Multiplication Subroutine are referred to from the main program, so they must be included with it. These are Subroutines I, II and III respectively in Appendix I.

Some of the probabilities that must be computed are obviously very small, so it is necessary to include a card which permits underflows in the machine. This card varies with the various computer centers, so it is not included in the program. Other control cards may be needed at the beginning of the program, before each subroutine, before the data card and at the very end.

The values needed for the input data card are K_1, K_2, K_3, L_a, D_B, Δ, α, L_A/Δ, C_N'/Δ, C_O/Δ. The parameter values should follow consecutively on the card with no spaces in between. The number of spaces allotted for each parameter are listed in Table 2.3.1. The decimal point in the Floating Point numbers must be placed in one of the six spaces; the Fixed Point numbers have no decimals.

As stated previously, the program prints out the two concentrations nearest to an α confidence level. It is also of interest to specify a concentration and then find

Table 2.3.1 Arrangement of Data Card

Parameter	Number of Spaces	
K_1	6	Floating Point
K_2	6	Floating Point
K_3	6	Floating Point
L_a	6	Floating Point
D_B	6	Floating Point
Δ	6	Floating Point
α	6	Floating Point
L_A/Δ	4	Fixed Point
C_N'/Δ	4	Fixed Point
C_O/Δ	4	Fixed Point

the corresponding probability of being above or below it. However, for dissolved oxygen this calculation would only be applicable for times of travel near the bottom of the sag. The probability of being below the specified concentration for small and large times of travel would be zero. For pollution, if one specified a concentration and determined the probability of being above it, then the probability for small travel times would decrease to zero.

Thus it was decided to find the concentration corresponding to a specified probability level. The probability of being above or below a certain concentration is easily obtained by summing probability mass densities.

III. OTHER INITIAL CONDITIONS

3.1 Variable Stream Conditions

Suppose it is assumed that upstream from t=0 for a reasonable distance there is no major source of pollution, that is, the river has fully recovered from any large sources of pollution upstream. The pollution that is flowing past the initial point is small and can be attributed to L_a. The dissolved oxygen is below its maximum due to the effects of D_B and L_a. At t=0 there may be some pollution entering. To describe the profile it is assumed that the pollution is entering at a constant rate. This case is included in order to cover the situation in which it is of interest to raise the level of pollution in the river to find out how much pollution the river can assimilate before the probability that the dissolved oxygen falls below the minimum levels exceeds certain preassigned values. This would be applicable in designing a treatment facility for an industrial plant or a municipality.

The initial probability generating function satisfying the above conditions will be the joint probability generating function in (2.2.3) evaluated at $t=\infty$ with the mean pollution state raised by a constant amount L_P/Δ. From (2.2.3)

$$P(S,R;\infty) = \exp\left[\left(\frac{L_a K_1}{\Delta K_2(K_1+K_3)} + \frac{D_B}{\Delta K}\right)\left(\frac{1-R}{R}\right) - \frac{L_a}{\Delta(K_1+K_3)}(1-S)\right]$$

$$\cdot\, R^{\frac{C_{N'}}{\Delta}} \quad . \tag{3.1.1}$$

This is the generating function for the condition in which only minor factors are involved. There are no residual effects of large sources of pollution.

Therefore, the initial probability generating function for this case is

$$P(S,R:0)=S^{\frac{L_B}{\Delta}} \quad \exp \quad [(\frac{L_a K_1}{K_2(K_1+K_3)} + \frac{D_B}{\Delta K_2}) \; [(\frac{1-R}{R}) -$$

$$\frac{L_a}{\Delta(K_1+K_3)} \; (1-S)] \qquad R^{\frac{C_{N'}}{\Delta}} \qquad (3.1.2)$$

that is, the right-hand side of (3.1.1) is multiplied by $S^{L_B/\Delta}$.

All of the assumptions and derivations of Section 2.1 are applicable for this case. The derivation here begins with the recommended procedure at the end of that section.

First, the substitution in step 1 at the end of Section 2.1 is made in *3.1.2). The result becomes

$$P(1-B- \frac{K_1}{K_2-(K_1+K_3)} \cdot \frac{1}{C'} \frac{C}{1+C}; \; 0) = (1-B- \frac{K_1}{K_2-(K_1+K_3)} \cdot \frac{1}{C}) \frac{L_B}{\Delta}$$

$$\exp \; [(\frac{L_a K_1}{\Delta K_2 (K_1+K3)} + \frac{D_B}{\Delta K_2}) \; (\frac{1}{C}) - \frac{L_a}{\Delta(K_1+K_3)}$$

$$(B + \frac{K_1}{K_2-(K_1+K_3)} \cdot \frac{1}{C})] \quad / \quad [(\frac{C}{1+C}) \frac{C_{N'}}{\Delta} \quad .$$

If the above is substituted in (2.1.12) and A,B, and C are replaced by expressions given in (2.1.7), (2.1.8) and (2.1.9) respectively, the result simplifies to

$$P(S,R;t)=\exp \quad [(\frac{L_a K_1}{\Delta K_2(K_1+K_3)} + \frac{D_B}{\Delta K_2}) \; (\frac{1-R}{R}) - \frac{L_a}{\Delta(K_1+K_3)}(1-S)]$$

$$\cdot R^{\frac{C_{N'}}{\Delta}} \quad (1-(1-S)e^{-(K_1+K_3)t} + \frac{K_1}{K_2-(K_1+K_3)} \; (e^{-(K_1+K_3)t}-e^{-K_2 t})$$

$$(\frac{1-R}{R})) \; ^{L_B/\Delta} \quad . \qquad (3.1 \qquad\qquad (3.1.3)$$

This is the joint probability generating function for both pollution and dissolved oxygen.

Allowing R to equal one in (3.1.3), one obtains the marginal probability generating function for pollution

$$P(S,1;t) = \exp\left[\frac{-L_a}{\Delta(K_1+K_3)}(1-S)\right] \cdot \left[1-(1-S)e^{-(K_1+K_3)t}\right]^{\frac{L_B}{\Delta}} . \qquad (3.1.4)$$

Expanding this in a power series in S, one obtains as a coefficient of S^M the probability of being in state M at time t,

$$P_M(t)=e^{\frac{-L_a}{\Delta(K_1+K_3)}} \sum_{a=0}^{M^*} \frac{L_B}{\Delta} \overset{-}{a} \quad \frac{L_a}{\Delta(K_1+K_3)} \quad (1-e^{-(K_1+K_3)t})^{\frac{L_B}{\Delta}}$$

$$- \frac{L_a}{\Delta(K_1+K_3)} -a \; (e^{-(K_1+K_3)t})^a \frac{\left(\frac{L_a}{(K_1+K_3)}\right)^{M-a}}{(M-a)!} . \qquad (3.1.5)$$

The result (3.1.5) could have been obtained from a different point of view, namely, using the theory of conditional distributions. Equation (2.2.5) which gives the probability mass densities for fixed initial conditions can be regarded as the probability density at time t conditioned on a fixed initial state. The probability density for the initial state of this section is formula (3.1.2). Forming the product of the conditional and marginal probability densities gives the joint probability that the pollution is in state M and the initial state is, say d. Then summing over the range of d, one obtains equation (3.1.5).

As explained in Appendix II, the moment generating function is found by replacing S in (3.1.4) by e^Θ, obtaining

$$M(\Theta,t)=\exp\left[\frac{-L_a}{\Delta(K_1+K_3)}(1-e^\Theta)\right] \cdot \left[1-(1-e^\Theta)e^{-(K_1+K_3)t}\right]^{\frac{L_B}{\Delta}} . \qquad (3.1.6)$$

From (3.1.6) the mean and variance of the amount of pollution present can be found to be

$$\text{Mean} = \frac{L_a}{K_1+K_3} + L_B\, e^{-(K_1+K_3)t} \qquad (3.1.7)$$

$$\text{Variance} = \Delta\left[\frac{L_a}{K_1+K_3} + L_B\, e^{-(K_1+K_3)t}(1-e^{-(K_1+K_3)t})\right]. \tag{3.1.8}$$

The mean is identical to the value of Dobbins (3), equation (1.2.2), if L_B is replaced by its equivalent $L_A - \frac{L_a}{K_1+K_3}$. Recall that L_B is defined as an increase in pollution, that is, the addition of pollution to the mean level of the river which is $\frac{L_a}{K_1+K_3}$ and L_A is defined as the total pollution. The variance consists of two terms, the first of which is the inherent river variability. The second term is the variability due to the increased pollution which is zero at the beginning of the reach. Then as t increases this term increases to a maximum and then decreases.

The marginal probability distribution for dissolved oxygen is found in a manner similar to that for pollution. If S is set equal to one in (3.1.3), then

$$P(1,R;t) = \exp\left[\left(\frac{L_a}{\Delta K_2(K_1+K_3)} + \frac{D_B}{K_2}\right)\left(\frac{1-R}{R}\right)\right] \cdot R^{\frac{C_N'}{\Delta}} .$$

$$\left[1 + \frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t} - e^{-K_2t})\left(\frac{1-R}{R}\right)\right]^{\frac{L_B}{\Delta}} . \tag{3.1.9}$$

As stated previously this can be expanded in a power series where the coefficient of R^N is the probability that the dissolved oxygen is in state N at time t. Negative values of N are not considered. Thus

$$P_N(t) = \exp\left[\frac{-L_aK_1}{\Delta K_2(K_1+K_3)} - \frac{D_B}{\Delta K_2}\right] \cdot \sum_{i=0}^{N^*} \left(1 - \frac{K_1}{K_2-(K_1+K_3)}\right.$$

$$\left.(e^{-(K_1+K_3)t} - e^{-K_2t})\right)^{\frac{L_B}{\Delta} - i} \cdot \left(\frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}\right.$$

$$\left. - e^{-K_2t})\right)^i \cdot \left[\frac{L_aK_1}{\Delta K_2(K_1+K_3)} + \frac{D_B}{\Delta K_2}\right]^{\frac{C_N'}{\Delta} - N - i} \cdot \frac{1}{\left(\frac{C_N'}{\Delta}-N-i\right)!} \tag{3.1.10}$$

$$\text{where } N^* = \begin{cases} \frac{C_{N'}}{\Delta} - N & \text{for } \frac{C_{N'}}{\Delta} - N \leq \frac{L_B}{\Delta} \\ \frac{L_B}{\Delta} & \text{for } \frac{C_{N'}}{\Delta} - N > \frac{L_B}{\Delta} . \end{cases}$$

Replace R by e^{ϕ} to get the moment generating function

$$M(\phi,t)=\exp\left[\left(\frac{L_a K_1}{\Delta K_2(K_1+K_3)} + \frac{D_B}{K_2}\right)(e^{-\phi}-1)\right] . e^{\phi} \frac{C_{N'}}{\Delta} .$$

$$\left[1 + \frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t} - e^{-K_2 t})(e^{-\phi}-1)\right] \frac{L_B}{\Delta} . \tag{3.1.11}$$

From this the mean and variance of the amount of dissolved oxygen can be found

$$\text{Mean} = C_{N'} - \frac{L_a K_1}{K_2(K_1+K_3)} - \frac{D_B}{K_2} - L_B \frac{K_1}{K_2-(K_1+K_3)}$$

$$(e^{-(K_1+K_3)t} - e^{-K_2 t}) \tag{3.1.12}$$

$$\text{Variance} = \Delta\left[\frac{L_a K_1}{K_2(K_1+K_3)} + \frac{D_B}{K_2} + L_B \frac{K_1}{K_2-(K_1+K_3)}\right.$$

$$\left.(e^{-(K_1+K_3)t} - e^{-K_2 t}).(1-(e^{-(K_1+K_3)t} - e^{-K_2 t}))\right] . \tag{3.1.13}$$

The mean in this form shows that for any time t the dissolved oxygen is at its maximum less an amount due to the factors that use oxygen. The variance is, of course, the sum of the factors that contribute to the variance. The mean value again checks with Dobbins' (3) result, equation (1.4), if L_B is replaced by $L_A - \frac{L_a}{K_1+K_3}$ and C_0 is replaced by $C_{N'} - \frac{L_a K_1}{K_2(K_1+K_3)} - \frac{D_B}{K_2}$. This second relation is true because it is the result when t is set equal to zero in (3.1.12).

A computer program was written for the formulas developed in this chapter. The explanation in Section 2.3 concerning that computer program (denoted as Program I) can serve as the explanation for this program. The print-outs are the same for both programs. The main program for this chapter is Program II in Appendix I. The Binomial Subroutine, the Exponential Subroutine, and the Series Multiplication Subroutine from Appendix I are needed also. The input data card for this program is almost exactly like the one for Program I, the exception being that for this data card the parameter $\frac{C_c}{\Delta}$ is not needed so that the spaces allotted to $\frac{C_c}{\Delta}$ are left blank. The format for this card is Table 2.3.1 also.

3.2 Measured Variability Initial Conditions

The problem to be studied in this chapter is, perhaps, the most frequently occurring one. It is assumed that the distribution at t=0 of both pollution and dissolved oxygen is a complex combination of the effects of various sources of upstream pollution and the pollution entering at t=0. Thus, no effort is made to find the distribution at t=0 by studying upstream effects on the river. The form of the distribution at t=0 must therefore be postulated and its mean and variance will be measured. If one assumes a random, non-systematic variability, a unimodal distribution seems to be a reasonable choice. A discrete distribution is needed and the binomial was chosen. The pollution and dissolved oxygen states have independent binomial distributions whose means and variances can be measured.

The pollution will have a mean state designated by L_A/Δ, the upper limit of the range will be R_2/Δ and the lower limit will be R_1/Δ. The parameter p (the probability of success in a Bernoulli trial) will then be $\frac{L_A-R_1}{R_2-R_1}$ and q is, as usual, 1-p. The probability generating function then will be

$$(pS+q)^{\frac{R_2-R_1}{\Delta}} \cdot S^{R_1/\Delta}$$

where the first term is the standard probability generating function for a binomial and the second term raises the mean level by R_1/Δ units from zero to R_1/Δ.

In a similar manner for the dissolved oxygen the mean state is designated by C_o/Δ and the range of states will be from a maximum of D_2/Δ to a minimum of D_1/Δ. The binomial parameter, denoted in this case as p', will be $\frac{C_0-D_1}{D_2-D_1}$ and q' is then 1-p'. The probability generating function will be

$$(p'R+q')^{\frac{D_2-D_1}{\Delta}} \cdot R^{\frac{D_1}{\Delta}}$$

where again the first factor is the binomial generating function with the lowest state zero and the second term raises each state by D_1/Δ states. Because of the independence assumption the joint distribution at t=0 will be the product of the two binomials, i.e.,

$$P(S,R;0)=(pS+q)^{\frac{R_2-R_1}{\Delta}} \cdot S^{R_1/\Delta} \cdot (p'R+q')^{(D_2-D_1)/\Delta} \cdot R^{D_1/\Delta} \tag{3.2.1}$$

All of the assumptions and derivations of Section 2.1 are applicable for this case. Mathematical manipulations of probability generating functions and moment generating functions can be found in Appendix II. The derivation here begins with the equations at the end of that section.

If we make the substitution of Step 1 of Section 2.1 into (3.2.1) we obtain

$$P(1-B-\frac{K_1}{K_2-(K_1+K_3)} \cdot \frac{1}{C}, \frac{C}{1+C}; 0) = [p(1-B-\frac{K_1}{K_2-(K_1+K_3)}$$

$$\cdot \frac{1}{C}) + q]^{\frac{R_2-R_1}{\Delta}} \cdot (1-B-\frac{K_1}{K_2-(K_1+K_3)} \cdot \frac{1}{C})^{R_1/\Delta}$$

$$\cdot (p'\frac{C}{1+C} + q')^{(D_2-D_1)/\Delta} \cdot (\frac{C}{1+C})^{D_1/\Delta}.$$

If we replace A, B, and C by their equals (as stated in Step 2) and simplify, the result becomes

$$P(S,R;t)=\exp\{[\frac{L_aK_1}{\Delta(K_1+K_3)(K_2-(K_1+K_3)}(1-e^{-(K_1+K_3)t})$$

$$-\frac{L_aK_1}{\Delta K_2[K_2-(K_1+K_3)]}(1-e^{-K_2t})+\frac{D_B}{\Delta K_2}(1-e^{-K_2t})][\frac{1-R}{R}]$$

$$-\frac{L_a}{\Delta(K_1+K_3)}(1-e^{-(K_1+K_3)t})(1-S)\}.$$

$$\frac{C_{N'}}{R^{\Delta}}(1+(\frac{1-R}{R})e^{-K_2t})^{(N'-D_2)/\Delta}\cdot[1+\frac{K_1}{K_2-(K_1+K_3)}$$

$$(e^{-(K_1+K_3)t}-e^{-K_2t})(\frac{1-R}{R})-(1-S)e^{-(K_1+K_3)t}]^{R_1/\Delta}.$$

$$[1+p\{\frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}-e^{-K_2t})(\frac{1-R}{R})-(1-S)$$

$$e^{-(K_1+K_3)t}\}]^{(R_2-R_1)/\Delta}.[1+q'e^{-K_2t}(\frac{1-R}{R})]^{(D_2-D_1)/\Delta}. \tag{3.2.3}$$

This is the joint probability generating function for both pollution and dissolved oxygen. The joint density $P_{M,N}(t)$ is the coefficient of S^MR^N in the power series expansion.

The marginal probability generating function for pollution is found by allowing R to be one in (3.2.3)

$$P(S,1;t)=\exp[-\frac{L_a}{\Delta(K_1+K_3)}(1-e^{-(K_1+K_3)t})(1-S)]\ .$$

$$[1-(1-S)e^{-(K_1+K_3)t}]^{R_1/\Delta}\ .\ 1-p(1-S)e^{-(K_1+K_3)t}]^{\frac{(R_2-R_1)}{\Delta}} \tag{3.2.4}$$

Expanding this in a power series in S, one obtains as a coefficient of S^M the probability of being in state M at time t.

The moment generating function is obtained from (3.2.4) by replacing S by e^{θ}

$$M(\theta,t)=\exp\left[-\frac{L_a}{\Delta(K_1+K_3)}(1-e^{-(K_1+K_3)t})(1-e^{\theta})\right].$$

$$[1-(1-e^{\theta})e^{-(K_1+K_3)t}]^{R_1/\Delta}\,.\,[1-p(1-e^{\theta})e^{-(K_1+K_3)t}]^{(R_2-R_1)/\Delta}\,. \tag{3.2.5}$$

The mean amount of pollution is once again found from (3.2.5). The definition of p was used in the simplification. The mean amount of pollution present is

$$\text{Mean} = \frac{L_a}{K_1+K_3}(1-e^{-(K_1+K_3)t})+L_A e^{-(K_1+K_3)t}. \tag{3.2.6}$$

This result is identical to the mean obtained by Dobbins (3), equation (1.2.2). The second term dominates for small values of t and the first term dominates for the larger values of t.

The variance of the amount of pollution present is also found from (3.2.5)

$$\text{Variance} = \Delta\Big[\frac{L_a}{K_1+K_3}(1-e^{-(K_1+K_3)t})+ R_1e^{-(K_1+K_3)t}.$$

$$(1-e^{-(K_1+K_3)t}) + (R_2-R_1)e^{-(K_1+K_3)t}.$$

$$(1-pe^{-(K_1+K_3)t})\Big]. \tag{3.2.7}$$

For $R_1 = R_2 = L_A$ this case reduces to that studied in Chapter II.

For the marginal probability generating function for dissolved oxygen we set S equal to unity in (3.2.3), obtaining

$$P(1,R;t) = \exp\Big\{\Big[\frac{L_aK_1}{\Delta(K_1+K_3)[K_2-(K_1+K_3)]}(1-e^{-(K_1+K_3)t})$$

$$-\frac{L_aK_1}{\Delta K_2[K_2-(K_1+K_3)]}(1-e^{-K_2t}) + \frac{D_B}{\Delta K_2}(1-e^{-K_2t})\Big]\Big[\frac{1-R}{R}\Big].$$

$$\frac{C_{N'}}{R\Delta}\left(1+\left(\frac{1-R}{R}\right)e^{-K_2t}\right)^{\frac{C_{N'}-D_2}{\Delta}}\,.\,\Big[1+\frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}$$

$$-e^{-K_2t}) \cdot (\frac{1-R}{R})]\ ^{R}1/\Delta \quad [1+p \frac{K_1}{K_2-(K_1+K_3)} \quad (e^{-(K_1+K_3)t} - e^{-K_2t})$$

$$\cdot (\frac{1-R}{R})] \quad \frac{R_2-R_1}{\Delta} \qquad 1+q' e^{-K_2t} \quad (\frac{1-R}{R})] \quad \frac{D_2-D_1}{\Delta} \cdot \qquad (3.2.8)$$

As stated previously this can be expanded in a power series where the coefficient of R^N is the probability that the dissolved oxygen is in state N at time t. Negative values of N are once again not considered.

For the moment generating function for dissolved oxygen replace R by e in (4.7) obtaining

$$M(\ ,t)=\exp \frac{L_aK_1}{\Delta(K_1+K_3) \quad [K_2-(K_1+K_3)]} \quad (1-e^{-(K_1+K_3)t})$$

$$- \frac{K_aK_1}{\Delta K_2[K_2-(K_1+K_3)]} \quad (1-e^{-K_2t}) + \frac{D_B}{\Delta K_2} \quad (1-e^{-K_2t})] \quad [e^{-\phi-1}]$$

$$\cdot\ e^{\phi \underline{C_{N'}}} \quad [1+ \quad (e^{-\phi}1) \ e^{-J_2t}] \ \frac{C_{N'}-D_2}{\Delta} \quad \cdot \ [1 + \frac{K_1}{K_2-(K_1+K_3)}$$

$$(e^{-(K_1+K_3)t} - e^{-K_2t})] \quad (e^{-\phi}-1) \quad R_{1/\Delta} \ [1+p \frac{K_1}{K_2-(K_1+K_3)}$$

$$(e^{-(K\ +K\ )t} - e^{-K_2t}) \quad (e^{-\phi}-1)] \qquad \frac{R_2-R_1}{\Delta} \qquad [1+\sigma' e^{-K_2t(e-\phi-1)}]$$

$$\frac{D_2-D_1}{\Delta}. \qquad (3.2.8)$$

From this the mean and variance of the amount of dissolved oxygen are found as follows

$$\text{Mean} = D_{N'} + \frac{L_aK_1}{K_2[K_2-(K_1+K_3)]} \ (1-e^{-K_2t}) - \frac{L_aK_1}{(K_1+K_3) \quad [K_2-(K_1+K_3)]}$$

$$(1-e^{-(K_1+K_3)t}) \ \frac{D_B}{K_2} \quad (1-e^{-K_2t}) - L_A \frac{K_1}{K_2-(K_1+K_3)} \qquad (3.2.9)$$

$$(e^{-(K_1+\ K_3)\ t} - e^{-K_2t}) \quad - \ (C_{N'}-C_0)e^{-K_2t}.$$

$$
\begin{aligned}
\text{Variance} = \Delta\Big[& \frac{L_a K_1}{(K_1+K_3)[K_2-(K_1+K_3)]}(1-e^{-(K_1+K_3)t}) \\
& - \frac{L_a K_1}{K_2[K_2-(K_1+K_3)]}(1-e^{-K_2 t}) + \frac{D_B}{K_2}(1-e^{-K_2 t}) \\
& + (C_{N'}-D_2)e^{-K_2 t}(1-e^{-K_2 t}) + \{R_1 \frac{K_1}{K_2-(K_1+K_3)} \\
& (e^{-(K_1+K_3)t}-e^{-K_2 t}).(1-\frac{K_1}{K_2-(K_1+K_3)} \\
& (e^{-(K_1+K_3)t}-e^{-K_2 t}))\} + \{(R_2-R_1)p \\
& \frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}-e^{-K_2 t}).(1-p\frac{K_1}{K_2-(K_1+K_3)} \\
& (e^{-(K_1+K_3)t}-e^{-K_2 t}))\} + (D_2-D_1)q'e^{-K_2 t} \\
& (1-q'e^{-K_2 t})\Big]. \qquad (3.2.10)
\end{aligned}
$$

The mean is identically equal to that of Dobbins (3), equation (1.2.4). The terms involving L_A and C_0 are dominant for small t. As t increases these terms are "damped out" and the others become dominant. Similarly for the variance, as t increases the importance of the initial conditions is less. For $R_1=R_2=L_A$ and $D_1=D_2=C_0$, this case reduces to the one originally studied in Chapter II.

Consider next the estimates for the values of L_A, R_1, and R_2 in the pollution initial condition. From a random sample one can get estimates for the mean and variance. This would enable two of the three to be estimated, but not the third one. The recommended procedure is as follows: estimate L_A from the observations. Then choose in a reasonable manner upper and lower limits R_2 and R_1. Perhaps the maximum and minimum values in a random sample would be reasonable estimates or better estimates might be $L_A \pm 3s$ where s is a sample standard deviation. One obviously outlying value should not determine either R_2 or R_1. In a similar manner estimates of C_0, D_1 and D_2 are obtained. The choice of values used for the ranges is not critical, in most cases, because with increasing time their effect is diminished. The initial mean values do, however, contribute significantly to the variance.

For this case, as with the previous cases, a computer program was written to perform the arithmetic involved in the use of the formulas. The explanation in section 2.3 is a satisfactory explanation of this program also. The printouts for the two programs are the same.

The main program for this chapter is Program III in Appendix I. The three subroutines I, II and III in Appendix I must be included when using the main program.

Because of the additional initial condition parameters the input data card for this case must be expanded. The parameters that must be specified are K_1, K_2, K_3, L_a, D_B, Δ, α, $\frac{L_A}{\Delta}$, $\frac{C_N'}{\Delta}$, $\frac{C_0}{\Delta}$, $\frac{R_2}{\Delta}$, $\frac{R_1}{\Delta}$, $\frac{D_2}{\Delta}$, and $\frac{D_1}{\Delta}$. The values follow consecutively on the card with no extra spaces in between the parameters. The number of spaces allotted to each parameter appears in Table 3.2.1.

Table 3.2.1 Arrangement of Data Card

Parameter	Number of Spaces	
K_1	6	Floating Point
K_2	6	Floating Point
K_3	6	Floating Point
L_a	6	Floating Point
D_B	6	Floating Point
Δ	6	Floating Point
α	6	Floating Point
L_A/Δ	4	Fixed Point
C_N'/Δ	4	Fixed Point
C_0/Δ	4	Fixed Point
R_2/Δ	4	Fixed Point
R_1/Δ	4	Fixed Point
D_2/Δ	4	Fixed Point
D_1/Δ	4	Fixed Point

IV. APPLICATIONS TO SACRAMENTO RIVER

4.1 Description

The background and data for this chapter come from Bulletin No. 111 of The Resources Agency of California, Department of Water Resources (10). One of the purposes of the report is to describe the conditions of the river with regard to organic pollution. The river is in good condition to support fish life but the amount of pollution is increasing. One reason that the quality must be maintained is that five cities use the river for their domestic water supply. Also, the river supports a large commercial fishing industry. The king salmon is hatched upstream, journeys to the ocean where after four years it returns to spawn and die. Any single grossly polluted reach could disrupt this process. Recreational use of the river is also important. Water contact sports are very popular so that the water quality affects the public health. Other large recreational opportunities are for boats and sport fishing.

The organic water pollution is increased by homes and towns and by the effluent of municipal sewage treatment plants. Industries also furnish organic pollution. Irrigation returns increase the mineral content but not the organic content of the river to any great extent.

The Sacramento River begins in northern California, flows southward past Sacramento and westward to the Pacific Ocean. The quantity of flow is controlled by the Shasta Dam in northern California so that the sea water cannot intrude inland during periods of low rainfall. The upper portion of the river passes through mountainous country and there is a relatively small amount of pollution entering. The water is cold which allows high concentrations of dissolved oxygen and the percent saturation is above 90%. The middle portion has some pollution entering from towns and small cities, but there is no place where there is a significant source of pollution which causes a sag in the dissolved oxygen. The pollution from towns is rapidly assimilated and the river rapidly recovers. This reach has many irrigation diversions and returns. Also, several rivers flow into the Sacramento in this reach. The water temperature continues to increase slowly which lowers the maximum solubility levels for dissolved oxygen. However, the percent saturation still remains high.

The lower portion, which is of special interest, is the region from Sacramento to the sea. There is very little irrigation water taken and there are no large tributaries. There is one large canal which takes water from the river, but it has no effect on the present analysis. The pollution from the city of Sacramento is large enough to cause a significant lowering of the oxygen levels below Sacramento. The water temperature remains relatively constant throughout the reach at about 70 degrees Fahrenheit. Hence it has no effect on the present analysis.

The three reaches were each sampled intensively at different seasons of the year, generally between May and October. An intensive survey lasts four days and during this time measurements of both pollution and dissolved oxygen were taken about eight times each day. Unfortunately many of the sampling stations measured only dissolved oxygen and not pollution; thus, the data available on pollution was limited. The actual measurements are included in the report.

The data in the report for the lower reach is confounded by tidal actions. The tidal action results in flow retardation throughout the reach. The Sacramento pollution enters at several points at approximately 50 miles above the mouth of the river and there are flow reversals due to tides up to approximately 44 miles above the mouth.

The report suggests that the data be censored so as to eliminate (as much as possible) the tidal effect. The method recommended in the report was adopted here and is explained in the following paragraph.

The time of low low water at various points was recorded as part of the intensive survey. Then by interpolation between these points the time of low low water can be found at each sampling station for each day. Then the times of slack water were determined by subtracting 30 minutes from the low low water times. The data used in this analysis is any data taken within one hour of slack water. Each sampling station thus contributes a maximum of four points, that is, one per day. Many times only three data points were available and sometimes only two or one, because no measurements were made at times near enough to slack water.

There are approximately thirty sampling stations which are used to give dissolved oxygen data and approximately three points per station, see Table 4.1.1. The situation with regard to pollution data is not good because as stated earlier very few of the 30 sampling stations recorded pollution measurements. With an average of three per station, one has only a few points with which to work. Also in this case, the initial concentration of pollution was not high enough to show a trend. Thus, minor fluctuations affected the shape of the curve. The report states that there is a large error in the analytical technique of measuring pollution when the pollution level is low. Therefore, because of the limited number of points and the large measurement variability of

Table 4.1.1 Sacramento River Data

River Mile	Time of Travel, days	Dissolved Oxygen, ppm	River Mile	Time of Travel, days	Dissolved Oxygen, ppm
50.8		8.5	28.4	1.33	7.4
		8.7			7.5
		8.7			7.6
49.8		8.5	27.4	1.52	7.3
		8.8			7.0
		8.9			7.6
48.4		8.6	26.8	1.63	7.1
		8.7			7.3
47.1		8.5			6.8
		8.7			7.4
		8.9	25.5	1.96	7.3
46.3	0.00	8.3			7.5
		8.6			6.8
		8.9	24.3	2.09	7.2
45.1	0.05	8.2			7.5
		8.7			6.8
		8.6			7.7
43.4	0.14	8.0	23.3	2.25	7.6
		8.6			7.1
		8.3			7.3
		8.4	22.3	2.43	7.2
42.1	0.22	8.1			7.3
		8.0			7.3
		8.4	21.1	2.65	7.2
		8.5			7.7
41.1	0.29	8.2			8.0
		8.0			7.5
		8.2	20.1	2.83	7.2
39.8	0.39	7.9			7.7
		7.9			8.2
		8.0			7.6
		8.3	18.8	3.03	7.2
38.6	0.47	7.7			7.8
		8.0			8.2
37.2	0.58	7.9	17.5	4.2	8.1
		7.9			8.5
35.9	0.68	7.8			8.7
34.4	0.80	8.1	15.1	5.6	8.1
33.5	0.87	8.0			8.4
		7.9			8.8
32.5	0.95	7.8	13.4	7.0	8.2
31.6	1.02	7.9			8.3
		7.7			8.6
		7.8			8.9
30.1	1.15	7.4	11.1	8.2	8.3
		7.5			8.4
		7.7			

the points, it was decided not to try to fit a pollution curve. The pollution data was used only as a guide to the value of L_A for the dissolved oxygen prediction.

The report shows plots of the dissolved oxygen for the slack following both high high water and low low water. It shows a definite oxygen sag. Then, in addition, the report makes the observation that the scatter of points tends to be greatest near the bottom of the sag and smaller near the upper levels. This is precisely the form of the variance function derived in earlier chapters. The magnitudes of the theoretical and actual variances will be examined later in this section.

4.2 Analysis of the Data

Before the estimates of the parameters are obtained it is necessary to decide which set of initial conditions are most applicable to these conditions. In the river above Sacramento there is only minor pollution from small sources along the river. In the area around the city of Sacramento a significant amount of pollution is added. This description is most closely associated with the assumptions of Chapter III, Section 3.1. Hence its results will be used here. The probability generating function for oxygen is

$$P(1,R;t) = \exp[(\frac{L_a K_1}{\Delta K_2(K_1+K_3)} + \frac{D_B}{\Delta K_2})(\frac{1-R}{R})] \,.R^{\frac{C_N'}{\Delta}}$$

$$[1 + \frac{K_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t} - e^{-K_2 t})(\frac{1-R}{R})]^{L_A/\Delta}$$

and the mean and variance for the amount of dissolved oxygen are

$$\text{Mean} = C_N' - \frac{L_a K_1}{K_2(K_1+K_3)} - \frac{D_B}{K_2} - L_A \cdot \frac{K_1}{K_2-(K_1+K_3)} \cdot$$

$$(e^{-(K_1+K_3)t} - e^{-K_2 t})$$

$$\text{Variance} = {}_{\Delta}\Big[\frac{L_a K_1}{K_2(K_1+K_3)} + \frac{D_B}{K_2} + L_A \cdot \frac{K_1}{K_2-(K_1+K_3)} \cdot$$

$$(e^{-(K_1+K_3)t} - e^{-K_2 t}) \; [1 - (e^{-(K_1+K_3)t} - e^{-K_2 t})]\Big] \; .$$

In order to apply the theory of Chapter III the times of travel between Sacramento and the sampling stations are needed. Time of travel t = 0 was chosen at Mile 46.3. This station is approximately where the oxygen begins to decrease. It is below the Sacramento Sewage Treatment Plant Outfalls. In this river, as in most rivers, the velocity is not constant so that the time of travel is not related to river mile by a simple equation. For a given flow of water the cross sectional area varies inversely with the velocity, see Fair and Geyer (6). Using these values and interpolating for intermediate points, one obtains the approximate times of travel between stations. The cumulative sums give the time of travel from Sacramento.

The times of travel values are listed in Table 4.1.1. The four or five stations nearest the ocean have very small velocities of 0.03 feet per second, so that the time of travel figures may be less accurate than they are at the upper stations. An error in these velocities might change the times of travel considerably. The effect would be to make two neighboring stations either closer together or further apart. However, the recovery is practically complete in this lower stretch, so that the dissolved oxygen levels show little change with time of travel.

Using only the factors of bacterial reaction with coefficient K_1 and reaeration with coefficient K_2, the report predicts a single value for the dissolved oxygen deficit at one point downstream. This prediction is the extent of the use of these concepts in the report. The parameters used in the present analysis were arrived at in several ways. The values are as follows:

$K_1 = 0.35$ per day $D_B = 0.10$ ppm/day
$K_2 = 0.75$ per day $K_3 = 0.20$ per day
$L_a = 0.20$ ppm/day $L_A = 6.80$ ppm.

The value of K_1 is taken directly from the report. Samples were taken and the rate of reaction was measured in the laboratory. This value is the approximate average. The K_2 value was chosen from within the range of K_2 values reported. Many of the reported values were negative and therefore impossible. As would be expected for this type of river the K_2 value is small. The river is deep so the exposed surface area per unit volume is small. The river has a small velocity and thus the turbulence is small. The three small parameters K_3, D_B, and L_a were obtained from pertinent measurements of the report, from studying the predicted mean value of dissolved oxygen at the extremes of $t = 0$ and $t = \infty$ and from fitting the curve to the observed data. These values are small in this case, so that any error in these would not significantly affect the probability distribution. The value used for L_A was obtained from the BOD data.

The only unknown remaining to be estimated is the parameter Δ. Since the mean amount of dissolved oxygen is independent of Δ, it cannot be used for the estimate. Thus, the estimate will be obtained from the variance function. Using the parameters and any downstream point, one can obtain an estimate of Δ by equating the sample and predicted variance. Several estimates could be combined to get a pooled estimate. For the purpose here the estimate will be obtained from the variance function, upstream from t = 0. The data in Table 4.1.1 between Mile 50.8 and Mile 47.1 was used. The pooled sample variance is 0.0298. At t = 0 the predicted variance is 0.303Δ. Thus $\Delta = 0.10$. The downstream data was not used for this estimate because it will be used to compare the observed variance with the predicted variance. Thus the estimate is independent of the data which will be used to test the theory.

The maximum solubility of oxygen in water at 21°C at sea level is 9 ppm. These are the conditions under which the study was made. As a result $\Delta N' = 9.0$.

This completes the information needed for the probability generating function and hence, the mean and variance functions. At any point in time the probability distribution can be given which describes the condition of the river.

The data points for various times of travel are plotted in Figure 4.2.1 along with the predicted mean function and the lower 20% and 10% confidence curves. Some points are plotted to the left of t = 0 to show that the mean level is steady until the effect of the Sacramento pollution becomes apparent. The last three travel times were omitted from the figure, because if they were included the time scale would be too small for small travel times. These points are strongly influenced by tidal action and they are, therefore, the least reliable. Table 4.1.1 shows these points to be above 8 ppm. The predicted mean rises from 8.2 to about 8.5. Thus the fit is good in this recovery zone.

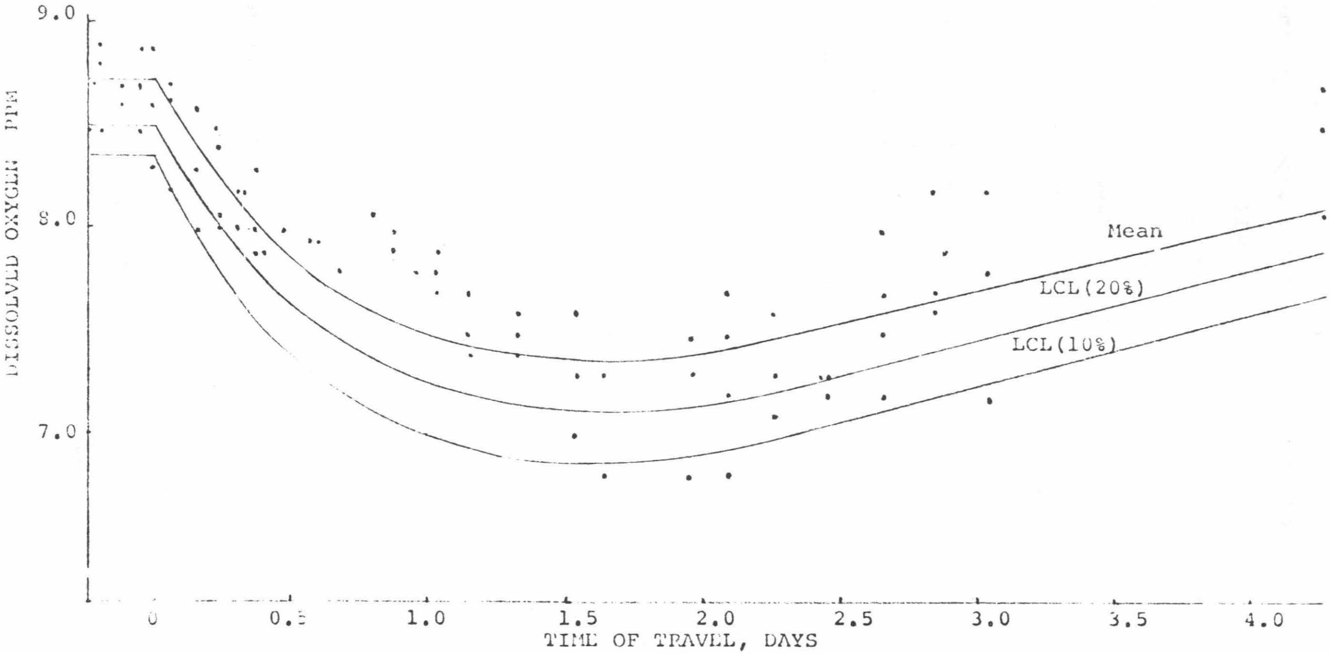

Figure 4.2.1. Mean and Lower Confidence Limits for Dissolved Oxygen

The predicted mean function is the upper curve in Figure 4.2.1. It fits the data reasonably well.

If the theoretical distribution is correct, then the lower 20 percent confidence curve should have roughly 20 percent of the data points below it. An actual count was made of the number of points below several such confidence bounds, with the results in Table 4.2.1.

Table 4.2.1. Number of Points Below Confidence Bounds.

Number of Points	Actual Percent of Points	Expected Percent of Points
17	18	20
8	8.6	10
0	0	1

The total number of data points is 93. It should also be noted that in Figure 4.2.1 the points below the confidence curve are spread out along the time axis and not confined to any one region.

One of the main uses of the results of this investigation will be for the determination of confidence bounds under future conditions. The agreement above between the actual percent and the predicted percent is excellent. There is reasonable evidence from the data used here that confidence bounds will give useful and accurate results.

There appears to be a slight jump in dissolved oxygen concentration in the region of "three-fourths day" time of travel. More points are needed in this region to make clear what is occurring. Nothing in the report gives reason to believe that there is a sudden increase in dissolved oxygen. For example, one might suspect that this is the confluence of the Sacramento River and another river of higher dissolved oxygen, but such is not the case.

Continuing with Figure 4.2.1 one notices that the confidence curves are farthest from the mean in the region of the bottom of the sag because the variance is larger than at the extremes. This is the most critical part of the curve because this is where the fish are most likely to be killed. Thus if a person added more pollution purely on the basis that the mean level would still allow a healthy fish population, fish kills could occur in the sag region because of the large variability about the mean.

Consider next the variability in the data. From the theory developed here, the variance for any time of travel can be predicted. The samples from each station provide estimates of the variance. However, the small sample sizes result in highly variable estimates, but the large number of sampling stations enables trends to be detected.

For the given parameters, Figure 4.2.2 shows a plot of the variance function curve for various times of travel. At $t = 0$ the variance equals 0.03 and this is also the limiting value of the variance for large t. The variance is greatest in the region of the oxygen sag which is the cause of the widened confidence limit previously discussed in Figure 4.2.1. Also plotted in Figure 4.2.2 are the sample variances from the data. Different markings are used to denote sample variances obtained from samples of size two, three and four. Even a sample of size four is small, so it is expected that the points in the figure will "bounce around" considerably. If one considers only the points where the sample size is four, the points are reasonably close to the curve. As stated when discussing Figure 4.2.1, the region around one day time of travel needs more data points so that the situation can be better understood. It appears

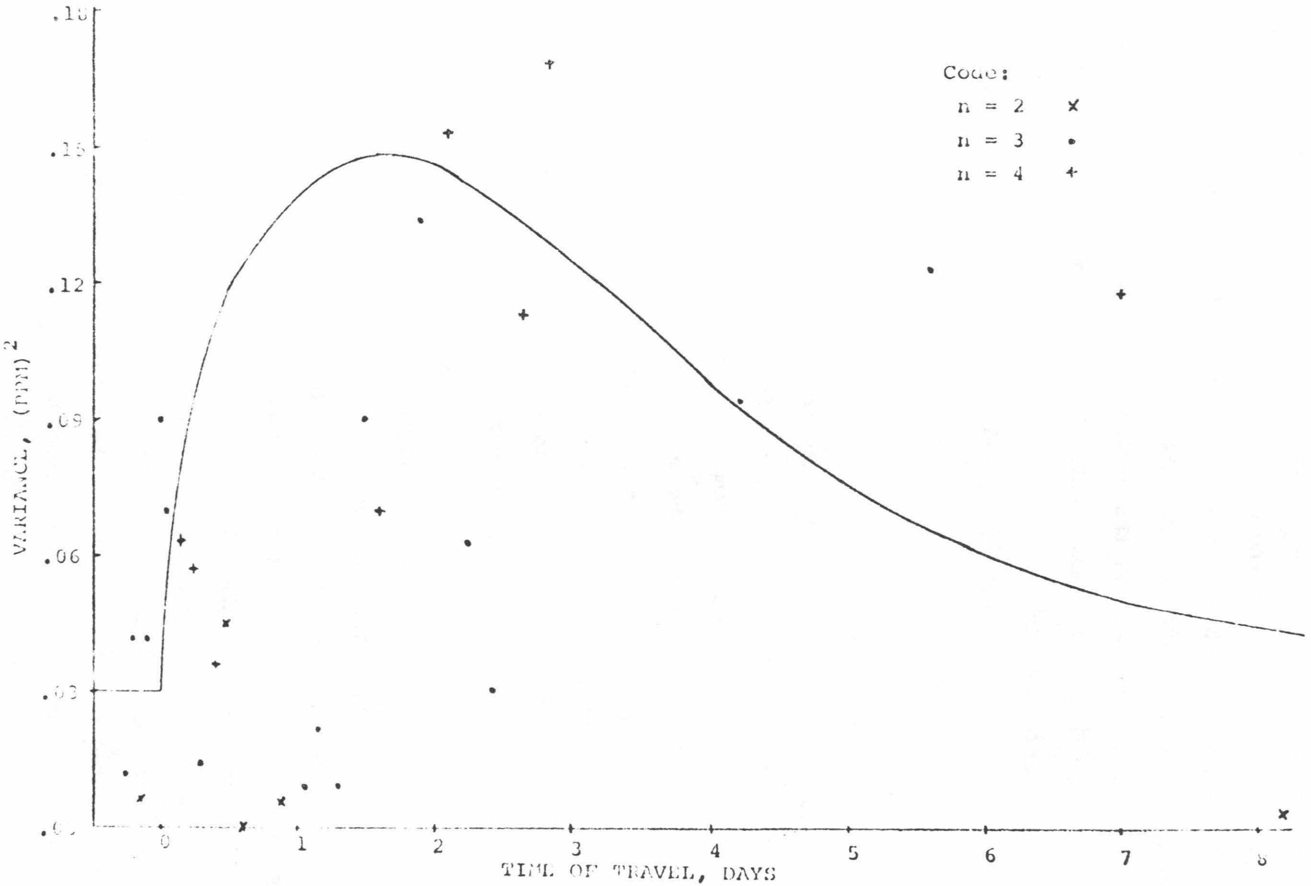

Figure 4.2.2. Predicted and Observed Variances

in Figure 4.2.2 that the within station variability is small and Figure 4.2.1 indicates a high between station variability. Also an investigation might reveal that stations do not have the same analytical errors in their measurements.

In conclusion it can be said that both the sample and predicted variances start at small values, increase to a maximum and then decrease.

With sample sizes less than five the range can often give a good indication of variability. Hence, it was decided to use a control chart for ranges to indicate the approximate expected ranges and 3σ limits. These are approximate because the theory of control charts is based on normal distribution theory. The data is from a non-normal population which has some skewness and a finite range. The sample range and the predicted mean range should be reasonably close, but one would not expect the larger ranges with truncated data that may be encountered with normally distributed data.

The predicted value for the mean range is found by multiplying the predicted σ by d_2 and the upper control limit is $D_2\sigma$. All lower control limits are zero. The multiplier d_2 is defined as the expected value of the range divided by σ and for normally distributed data d_2 is a function only of the sample size. The upper control limit or 3σ limit is $d_2\sigma + 3\sigma_R$ where σ_R is the standard deviation of the range. The variance σ_R equals σ times a function of the sample size for a normal population. Hence, the upper control limit is given by $D_2\sigma$ where D_2 is a function of the sample size only. The probability of having a value above the 3σ limit is 0.005. The lower 3σ limit is similarly defined, but for small sample sizes this limit is negative. Thus a zero value for the lower limit is chosen.

Figure 4.2.3 is a control chart for ranges with a known population variance. Note that the time scale is changed from the upper to the lower portion. In general, the sample points are near and on both sides of the expected range center line. None of the points are out of control which is due to the fact that the population is truncated, especially on the upper side. This should be clarified in that it is possible in daylight with photosynthesis to get concentrations of dissolved oxygen higher than 9 ppm, but it is known that photosynthesis is not occurring to such an extent in this river.

The points to the left of t = 0 have the same predicted variance as the points at t = 0. They are included in the control chart as an indication that the variability is at a constant level until the oxygen begins to sag from the city of Sacramento pollution.

The next point of interest in Figure 4.2.3 is the region in the vicinity of one day of travel. There is a run of several points below the predicted range. This is the region discussed previously in connection with Figure 4.2.1 of the dissolved oxygen mean level. Although the variability within stations appears to be small on the control chart, Figure 4.2.1 shows that there is considerable variability between stations. Perhaps more data in this region would show less abrupt changes in the mean level between stations and more variability within each sampling station. There are several stations which contribute only one data point and thus do not appear on the control chart. More data at these points would have been most helpful.

The data in the control chart show that the variability is large near the sag portion and smaller before the sag and after the river recovers. This is the conclusion reached theoretically and plotted in Figure 4.2.2.

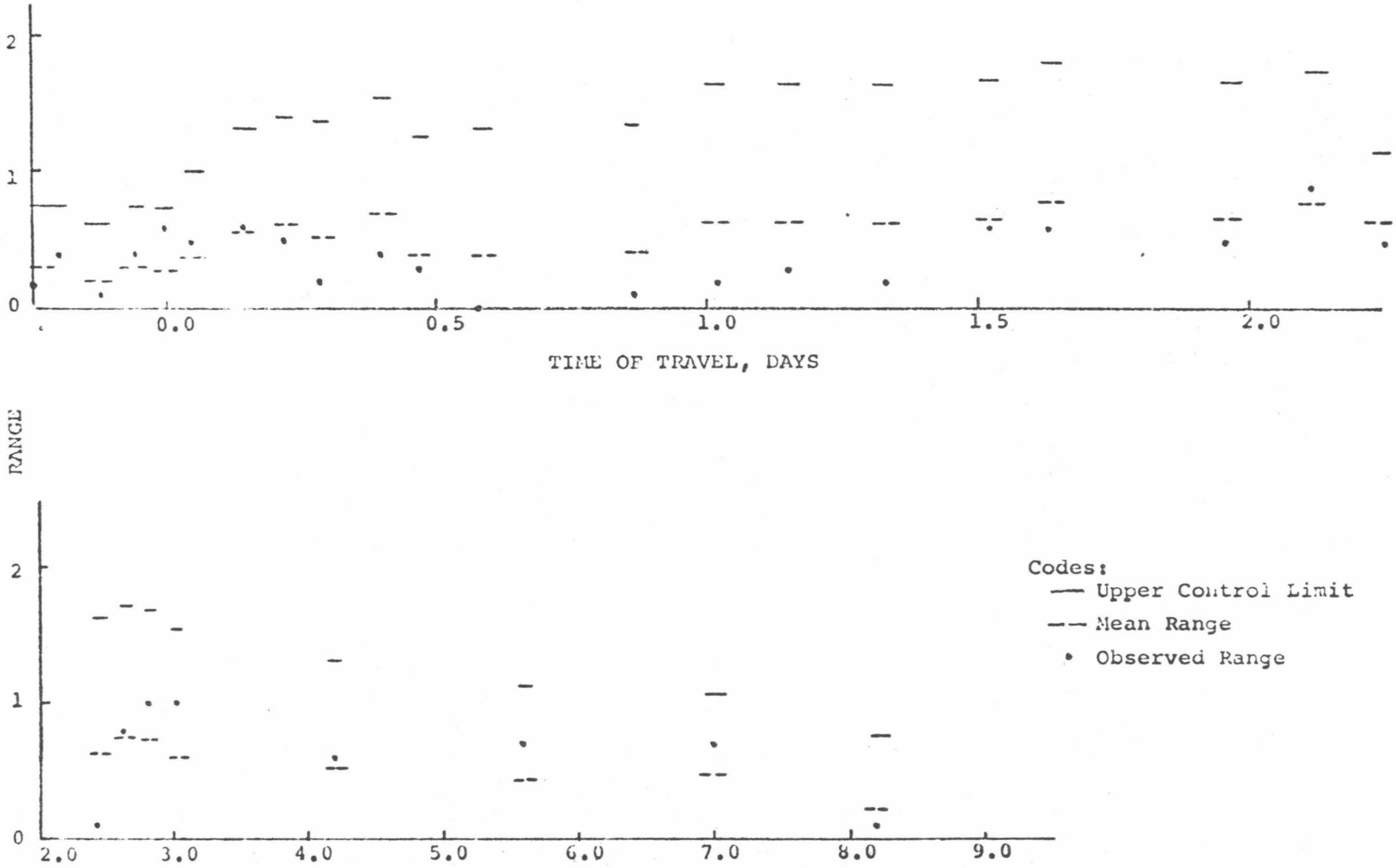

Figure 4.2.3. Control Chart for Ranges

This completes the analysis of the data. The theoretical results have been verified by this data as well as can be expected. More data is needed for parameter and time of travel estimation. It would be preferable to space the stations more evenly according to time of travel, that is, fewer stations at the higher velocities near Sacramento and more stations further downstream.

4.3 Applications to Future Situations

The theory developed in this dissertation has been used to describe the Sacramento River dissolved oxygen profile. The next question concerns future population and industrial growth and its affect on the river. With new sets of parameters the resulting new profile can be predicted. There are many possible changes that are reasonable to consider, and one such change will be considered here.

Suppose that the pollution from the city of Sacramento were to increase to a higher level, that is L_B, the increase over ordinary river amounts, was at a value of 15 ppm. Suppose also that pollution coming in from the sides, denoted by L_a, increased. The parameter L_a is defined as the change in level of the entire river by the pollution from the sides in one day. Because of the large flow or quantity of water in the river it takes a lot of pollution to raise its level. Hence it is reasonable that L_a was small in the last section and the value used now will be only $L_a = 0.50$. However, although it takes a lot of pollution to increase the river level, it takes a lot of change to improve the river condition when the sag becomes too great.

With these two changes in the parameters from Section 4.2, the parameters are now

$K_1 = 0.35$ per day $\qquad\qquad D_B = 0.10$ ppm/day

K_2 = 0.75 per day K_3 = 0.20 per day

L_a = 0.50 ppm/day L_B = 15 ppm.

These new parameters were used in the computer program and the new profile obtained. The complete printout is in Table 4.3.1. Looking at the probability mass densities first, it is evident at what concentrations the stream is most likely to be found. Also the probabilities of extreme values are given. By finding the cumulative probabilities one can find the concentration corresponding to a given cumulative probability and one can also find the probability of being below a given concentration. The shape of the distribution can be seen from a plot of the probabilities.

Figure 4.3.1 shows plots of the pollution and dissolved oxygen mass densities for a time of travel of two days. The pollution plot is unimodal and at this point is almost symmetrical. Other plots will show some skewness at the two extremes of travel times. The dissolved oxygen plot is also symmetrical, but other plots show some skewness at the extremes because of truncation. The lower scales are the same, so that the higher maximum of the dissolved oxygen shows that it has a smaller variance.

Figure 4.3.2 shows both the pollution and oxygen means plotted versus time of travel and also their corresponding 10 per cent confidence bounds. Note that the scales for the two concentrations are different, which makes the pollution confidence bound seem closer to the mean. The pollution mean is just a simple exponential decay and because of L_a it does not have zero as its lower limit. The dissolved oxygen profile is similar to the profile obtained in the previous section except that it starts lower and has a lower minimum.

Table 4.3.1. Computer Printout For Future Conditions

```
TIME OF TRAVEL IN DAYS = 1

PROBABILITY MASS DENSITIES FOR POLLUTION
PROBABILITY OF HAVING   7.00 PPM = 0.0000
PROBABILITY OF HAVING   7.10 PPM = 0.0001
PROBABILITY OF HAVING   7.20 PPM = 0.0001
PROBABILITY OF HAVING   7.30 PPM = 0.0002
PROBABILITY OF HAVING   7.40 PPM = 0.0003
PROBABILITY OF HAVING   7.50 PPM = 0.0005
PROBABILITY OF HAVING   7.60 PPM = 0.0008
PROBABILITY OF HAVING   7.70 PPM = 0.0013
PROBABILITY OF HAVING   7.80 PPM = 0.0019
PROBABILITY OF HAVING   7.90 PPM = 0.0028
PROBABILITY OF HAVING   8.00 PPM = 0.0040
PROBABILITY OF HAVING   8.10 PPM = 0.0056
PROBABILITY OF HAVING   8.20 PPM = 0.0077
PROBABILITY OF HAVING   8.30 PPM = 0.0103
PROBABILITY OF HAVING   8.40 PPM = 0.0134
PROBABILITY OF HAVING   8.50 PPM = 0.0172
PROBABILITY OF HAVING   8.60 PPM = 0.0215
PROBABILITY OF HAVING   8.70 PPM = 0.0262
PROBABILITY OF HAVING   8.80 PPM = 0.0313
PROBABILITY OF HAVING   8.90 PPM = 0.0366
PROBABILITY OF HAVING   9.00 PPM = 0.0419
PROBABILITY OF HAVING   9.10 PPM = 0.0468
PROBABILITY OF HAVING   9.20 PPM = 0.0512
PROBABILITY OF HAVING   9.30 PPM = 0.0548
PROBABILITY OF HAVING   9.40 PPM = 0.0574
PROBABILITY OF HAVING   9.50 PPM = 0.0588
PROBABILITY OF HAVING   9.60 PPM = 0.0589
PROBABILITY OF HAVING   9.70 PPM = 0.0577
PROBABILITY OF HAVING   9.80 PPM = 0.0554
PROBABILITY OF HAVING   9.90 PPM = 0.0520
PROBABILITY OF HAVING  10.00 PPM = 0.0478
PROBABILITY OF HAVING  10.10 PPM = 0.0429
PROBABILITY OF HAVING  10.20 PPM = 0.0377
PROBABILITY OF HAVING  10.30 PPM = 0.0325
PROBABILITY OF HAVING  10.40 PPM = 0.0274
PROBABILITY OF HAVING  10.50 PPM = 0.0226
PROBABILITY OF HAVING  10.60 PPM = 0.0182
PROBABILITY OF HAVING  10.70 PPM = 0.0144
PROBABILITY OF HAVING  10.80 PPM = 0.0111
PROBABILITY OF HAVING  10.90 PPM = 0.0084
PROBABILITY OF HAVING  11.00 PPM = 0.0062
PROBABILITY OF HAVING  11.10 PPM = 0.0045
PROBABILITY OF HAVING  11.20 PPM = 0.0032
PROBABILITY OF HAVING  11.30 PPM = 0.0022
PROBABILITY OF HAVING  11.40 PPM = 0.0015
PROBABILITY OF HAVING  11.50 PPM = 0.0010
PROBABILITY OF HAVING  11.60 PPM = 0.0006
PROBABILITY OF HAVING  11.70 PPM = 0.0004
PROBABILITY OF HAVING  11.80 PPM = 0.0003
PROBABILITY OF HAVING  11.90 PPM = 0.0002
PROBABILITY OF HAVING  12.00 PPM = 0.0001
PROBABILITY OF HAVING  12.10 PPM = 0.0001
PROBABILITY OF HAVING  12.20 PPM = 0.0000

MEAN = 9.56          VARIANCE = 0.4570
```

Table 4.3.1 Continued

```
PROBABILITY OF HAVING MORE THAN  10.40 PPM = 0.0949
PROBABILITY OF HAVING MORE THAN  10.30 PPM = 0.1223


TIME OF TRAVEL IN DAYS = 1
PROBABILITY MASS DENSITIES FOR OXYGEN
PROBABILITY OF HAVING     7.60 PPM = 0.0000
PROBABILITY OF HAVING     7.50 PPM = 0.0001
PROBABILITY OF HAVING     7.40 PPM = 0.0002
PROBABILITY OF HAVING     7.30 PPM = 0.0004
PROBABILITY OF HAVING     7.20 PPM = 0.0009
PROBABILITY OF HAVING     7.10 PPM = 0.0017
PROBABILITY OF HAVING     7.00 PPM = 0.0030
PROBABILITY OF HAVING     6.90 PPM = 0.0050
PROBABILITY OF HAVING     6.80 PPM = 0.0080
PROBABILITY OF HAVING     6.70 PPM = 0.0122
PROBABILITY OF HAVING     6.60 PPM = 0.0177
PROBABILITY OF HAVING     6.50 PPM = 0.0245
PROBABILITY OF HAVING     6.40 PPM = 0.0325
PROBABILITY OF HAVING     6.30 PPM = 0.0413
PROBABILITY OF HAVING     6.20 PPM = 0.0503
PROBABILITY OF HAVING     6.10 PPM = 0.0588
PROBABILITY OF HAVING     6.00 PPM = 0.0662
PROBABILITY OF HAVING     5.90 PPM = 0.0717
PROBABILITY OF HAVING     5.80 PPM = 0.0748
PROBABILITY OF HAVING     5.70 PPM = 0.0752
PROBABILITY OF HAVING     5.60 PPM = 0.0731
PROBABILITY OF HAVING     5.50 PPM = 0.0686
PROBABILITY OF HAVING     5.40 PPM = 0.0622
PROBABILITY OF HAVING     5.30 PPM = 0.0547
PROBABILITY OF HAVING     5.20 PPM = 0.0465
PROBABILITY OF HAVING     5.10 PPM = 0.0383
PROBABILITY OF HAVING     5.00 PPM = 0.0306
PROBABILITY OF HAVING     4.90 PPM = 0.0237
PROBABILITY OF HAVING     4.80 PPM = 0.0178
PROBABILITY OF HAVING     4.70 PPM = 0.0130
PROBABILITY OF HAVING     4.60 PPM = 0.0093
PROBABILITY OF HAVING     4.50 PPM = 0.0064
PROBABILITY OF HAVING     4.40 PPM = 0.0043
PROBABILITY OF HAVING     4.30 PPM = 0.0028
PROBABILITY OF HAVING     4.20 PPM = 0.0018
PROBABILITY OF HAVING     4.10 PPM = 0.0011
PROBABILITY OF HAVING     4.00 PPM = 0.0007
PROBABILITY OF HAVING     3.90 PPM = 0.0004
PROBABILITY OF HAVING     3.80 PPM = 0.0002
PROBABILITY OF HAVING     3.70 PPM = 0.0001
PROBABILITY OF HAVING     3.60 PPM = 0.0001
PROBABILITY OF HAVING     3.50 PPM = 0.0000

MEAN = 5.70     VARIANCE = 0.3016

PROBABILITY OF HAVING LESS THAN  5.10 PPM = 0.1124
PROBABILITY OF HAVING LESS THAN  5.00 PPM = 0.0818


TIME OF TRAVEL IN DAYS = 2

PROBABILITY MASS DENSITIES FOR POLLUTION
PROBABILITY OF HAVING     3.50 PPM = 0.0000
PROBABILITY OF HAVING     3.60 PPM = 0.0001
PROBABILITY OF HAVING     3.70 PPM = 0.0001
```

Table 4.3.1 Continued

```
PROBABILITY OF HAVING     3.80 PPM = 0.0002
PROBABILITY OF HAVING     3.90 PPM = 0.0004
PROBABILITY OF HAVING     4.00 PPM = 0.0007
PROBABILITY OF HAVING     4.10 PPM = 0.0011
PROBABILITY OF HAVING     4.20 PPM = 0.0017
PROBABILITY OF HAVING     4.30 PPM = 0.0027
PROBABILITY OF HAVING     4.40 PPM = 0.0040
PROBABILITY OF HAVING     4.50 PPM = 0.0058
PROBABILITY OF HAVING     4.60 PPM = 0.0081
PROBABILITY OF HAVING     4.70 PPM = 0.0111
PROBABILITY OF HAVING     4.80 PPM = 0.0147
PROBABILITY OF HAVING     4.90 PPM = 0.0197
PROBABILITY OF HAVING     5.00 PPM = 0.0239
PROBABILITY OF HAVING     5.10 PPM = 0.0294
PROBABILITY OF HAVING     5.20 PPM = 0.0351
PROBABILITY OF HAVING     5.30 PPM = 0.0410
PROBABILITY OF HAVING     5.40 PPM = 0.0466
PROBABILITY OF HAVING     5.50 PPM = 0.0516
PROBABILITY OF HAVING     5.60 PPM = 0.0558
PROBABILITY OF HAVING     5.70 PPM = 0.0590
PROBABILITY OF HAVING     5.80 PPM = 0.0608
PROBABILITY OF HAVING     5.90 PPM = 0.0612
PROBABILITY OF HAVING     6.00 PPM = 0.0602
PROBABILITY OF HAVING     6.10 PPM = 0.0578
PROBABILITY OF HAVING     6.20 PPM = 0.0543
PROBABILITY OF HAVING     6.30 PPM = 0.0498
PROBABILITY OF HAVING     6.40 PPM = 0.0447
PROBABILITY OF HAVING     6.50 PPM = 0.0393
PROBABILITY OF HAVING     6.60 PPM = 0.0337
PROBABILITY OF HAVING     6.70 PPM = 0.0283
PROBABILITY OF HAVING     6.80 PPM = 0.0233
PROBABILITY OF HAVING     6.90 PPM = 0.0187
PROBABILITY OF HAVING     7.00 PPM = 0.0148
PROBABILITY OF HAVING     7.10 PPM = 0.0114
PROBABILITY OF HAVING     7.20 PPM = 0.0086
PROBABILITY OF HAVING     7.30 PPM = 0.0064
PROBABILITY OF HAVING     7.40 PPM = 0.0046
PROBABILITY OF HAVING     7.50 PPM = 0.0033
PROBABILITY OF HAVING     7.60 PPM = 0.0023
PROBABILITY OF HAVING     7.70 PPM = 0.0016
PROBABILITY OF HAVING     7.80 PPM = 0.0010
PROBABILITY OF HAVING     7.90 PPM = 0.0007
PROBABILITY OF HAVING     8.00 PPM = 0.0004
PROBABILITY OF HAVING     8.10 PPM = 0.0003
PROBABILITY OF HAVING     8.20 PPM = 0.0002
PROBABILITY OF HAVING     8.30 PPM = 0.0001
PROBABILITY OF HAVING     8.40 PPM = 0.0001
PROBABILITY OF HAVING     8.50 PPM = 0.0000

MEAN = 5.90    VARIANCE = 0.4240

PROBABILITY OF HAVING MORE THAN  6.70 PPM = 0.0977
PROBABILITY OF HAVING MORE THAN  6.60 PPM = 0.1260

TIME OF TRAVEL IN DAYS = 2

PROBABILITY MASS DENSITIES FOR OXYGEN
PROBABILITY OF HAVING     7.50 PPM = 0.0000
PROBABILITY OF HAVING     7.40 PPM = 0.0001
PROBABILITY OF HAVING     7.30 PPM = 0.0002
```

Table 4.3.1 Continued

```
PROBABILITY OF HAVING    7.20 PPM = 0.0004
PROBABILITY OF HAVING    7.10 PPM = 0.0008
PROBABILITY OF HAVING    7.00 PPM = 0.0015
PROBABILITY OF HAVING    6.90 PPM = 0.0027
PROBABILITY OF HAVING    6.80 PPM = 0.0045
PROBABILITY OF HAVING    6.70 PPM = 0.0073
PROBABILITY OF HAVING    6.60 PPM = 0.0111
PROBABILITY OF HAVING    6.50 PPM = 0.0161
PROBABILITY OF HAVING    6.40 PPM = 0.0225
PROBABILITY OF HAVING    6.30 PPM = 0.0299
PROBABILITY OF HAVING    6.20 PPM = 0.0383
PROBABILITY OF HAVING    6.10 PPM = 0.0470
PROBABILITY OF HAVING    6.00 PPM = 0.0555
PROBABILITY OF HAVING    5.90 PPM = 0.0631
PROBABILITY OF HAVING    5.80 PPM = 0.0690
PROBABILITY OF HAVING    5.70 PPM = 0.0729
PROBABILITY OF HAVING    5.60 PPM = 0.0742
PROBABILITY OF HAVING    5.50 PPM = 0.0731
PROBABILITY OF HAVING    5.40 PPM = 0.0695
PROBABILITY OF HAVING    5.30 PPM = 0.0640
PROBABILITY OF HAVING    5.20 PPM = 0.0571
PROBABILITY OF HAVING    5.10 PPM = 0.0493
PROBABILITY OF HAVING    5.00 PPM = 0.0413
PROBABILITY OF HAVING    4.90 PPM = 0.0335
PROBABILITY OF HAVING    4.80 PPM = 0.0264
PROBABILITY OF HAVING    4.70 PPM = 0.0202
PROBABILITY OF HAVING    4.60 PPM = 0.0151
PROBABILITY OF HAVING    4.50 PPM = 0.0109
PROBABILITY OF HAVING    4.40 PPM = 0.0077
PROBABILITY OF HAVING    4.30 PPM = 0.0052
PROBABILITY OF HAVING    4.20 PPM = 0.0035
PROBABILITY OF HAVING    4.10 PPM = 0.0023
PROBABILITY OF HAVING    4.00 PPM = 0.0014
PROBABILITY OF HAVING    3.90 PPM = 0.0009
PROBABILITY OF HAVING    3.80 PPM = 0.0005
PROBABILITY OF HAVING    3.70 PPM = 0.0003
PROBABILITY OF HAVING    3.60 PPM = 0.0002
PROBABILITY OF HAVING    3.50 PPM = 0.0001
PROBABILITY OF HAVING    3.40 PPM = 0.0001
PROBABILITY OF HAVING    3.30 PPM = 0.0000

MEAN = 5.56   VARIANCE = 0.3122

PROBABILITY OF HAVING LESS THAN  5.00 PPM = 0.1284
PROBABILITY OF HAVING LESS THAN  4.90 PPM = 0.0949

TIME OF TRAVEL IN DAYS = 3

PROBABILITY MASS DENSITIES FOR POLLUTION
PROBABILITY OF HAVING    1.70 PPM = 0.0000
PROBABILITY OF HAVING    1.80 PPM = 0.0001
PROBABILITY OF HAVING    1.90 PPM = 0.0001
PROBABILITY OF HAVING    2.00 PPM = 0.0003
PROBABILITY OF HAVING    2.10 PPM = 0.0005
PROBABILITY OF HAVING    2.20 PPM = 0.0010
PROBABILITY OF HAVING    2.30 PPM = 0.0017
PROBABILITY OF HAVING    2.40 PPM = 0.0029
PROBABILITY OF HAVING    2.50 PPM = 0.0047
PROBABILITY OF HAVING    2.60 PPM = 0.0073
```

Table 4.3.1 Continued

```
PROBABILITY OF HAVING    2.70 PPM = 0.0109
PROBABILITY OF HAVING    2.80 PPM = 0.0154
PROBABILITY OF HAVING    2.90 PPM = 0.0211
PROBABILITY OF HAVING    3.00 PPM = 0.0277
PROBABILITY OF HAVING    3.10 PPM = 0.0351
PROBABILITY OF HAVING    3.20 PPM = 0.0428
PROBABILITY OF HAVING    3.30 PPM = 0.0505
PROBABILITY OF HAVING    3.40 PPM = 0.0575
PROBABILITY OF HAVING    3.50 PPM = 0.0634
PROBABILITY OF HAVING    3.60 PPM = 0.0676
PROBABILITY OF HAVING    3.70 PPM = 0.0698
PROBABILITY OF HAVING    3.80 PPM = 0.0699
PROBABILITY OF HAVING    3.90 PPM = 0.0679
PROBABILITY OF HAVING    4.00 PPM = 0.0640
PROBABILITY OF HAVING    4.10 PPM = 0.0586
PROBABILITY OF HAVING    4.20 PPM = 0.0521
PROBABILITY OF HAVING    4.30 PPM = 0.0450
PROBABILITY OF HAVING    4.40 PPM = 0.0379
PROBABILITY OF HAVING    4.50 PPM = 0.0310
PROBABILITY OF HAVING    4.60 PPM = 0.0248
PROBABILITY OF HAVING    4.70 PPM = 0.0192
PROBABILITY OF HAVING    4.80 PPM = 0.0146
PROBABILITY OF HAVING    4.90 PPM = 0.0108
PROBABILITY OF HAVING    5.00 PPM = 0.0078
PROBABILITY OF HAVING    5.10 PPM = 0.0055
PROBABILITY OF HAVING    5.20 PPM = 0.0037
PROBABILITY OF HAVING    5.30 PPM = 0.0025
PROBABILITY OF HAVING    5.40 PPM = 0.0016
PROBABILITY OF HAVING    5.50 PPM = 0.0011
PROBABILITY OF HAVING    5.60 PPM = 0.0007
PROBABILITY OF HAVING    5.70 PPM = 0.0004
PROBABILITY OF HAVING    5.80 PPM = 0.0002
PROBABILITY OF HAVING    5.90 PPM = 0.0001
PROBABILITY OF HAVING    6.00 PPM = 0.0001
PROBABILITY OF HAVING    6.10 PPM = 0.0000

MEAN = 3.79    VARIANCE = 0.3237

PROBABILITY OF HAVING MORE THAN  4.50 PPM = 0.0931
PROBABILITY OF HAVING MORE THAN  4.40 PPM = 0.1242

TIME OF TRAVEL IN DAYS = 3

PROBABILITY MASS DENSITIES FOR OXYGEN
PROBABILITY OF HAVING    8.00 PPM = 0.0000
PROBABILITY OF HAVING    7.90 PPM = 0.0001
PROBABILITY OF HAVING    7.80 PPM = 0.0002
PROBABILITY OF HAVING    7.70 PPM = 0.0004
PROBABILITY OF HAVING    7.60 PPM = 0.0008
PROBABILITY OF HAVING    7.50 PPM = 0.0016
PROBABILITY OF HAVING    7.40 PPM = 0.0029
PROBABILITY OF HAVING    7.30 PPM = 0.0052
PROBABILITY OF HAVING    7.20 PPM = 0.0087
PROBABILITY OF HAVING    7.10 PPM = 0.0136
PROBABILITY OF HAVING    7.00 PPM = 0.0201
PROBABILITY OF HAVING    6.90 PPM = 0.0282
PROBABILITY OF HAVING    6.80 PPM = 0.0376
PROBABILITY OF HAVING    6.70 PPM = 0.0478
PROBABILITY OF HAVING    6.60 PPM = 0.0579
```

Table 4.3.1 Continued

```
PROBABILITY OF HAVING     6.50 PPM = 0.0669
PROBABILITY OF HAVING     6.40 PPM = 0.0741
PROBABILITY OF HAVING     6.30 PPM = 0.0786
PROBABILITY OF HAVING     6.20 PPM = 0.0800
PROBABILITY OF HAVING     6.10 PPM = 0.0782
PROBABILITY OF HAVING     6.00 PPM = 0.0736
PROBABILITY OF HAVING     5.90 PPM = 0.0667
PROBABILITY OF HAVING     5.80 PPM = 0.0582
PROBABILITY OF HAVING     5.70 PPM = 0.0491
PROBABILITY OF HAVING     5.60 PPM = 0.0399
PROBABILITY OF HAVING     5.50 PPM = 0.0314
PROBABILITY OF HAVING     5.40 PPM = 0.0239
PROBABILITY OF HAVING     5.30 PPM = 0.0176
PROBABILITY OF HAVING     5.20 PPM = 0.0126
PROBABILITY OF HAVING     5.10 PPM = 0.0087
PROBABILITY OF HAVING     5.00 PPM = 0.0058
PROBABILITY OF HAVING     4.90 PPM = 0.0038
PROBABILITY OF HAVING     4.80 PPM = 0.0024
PROBABILITY OF HAVING     4.70 PPM = 0.0015
PROBABILITY OF HAVING     4.60 PPM = 0.0009
PROBABILITY OF HAVING     4.50 PPM = 0.0005
PROBABILITY OF HAVING     4.40 PPM = 0.0003
PROBABILITY OF HAVING     4.30 PPM = 0.0002
PROBABILITY OF HAVING     4.20 PPM = 0.0001
PROBABILITY OF HAVING     4.10 PPM = 0.0000

MEAN = 6.17     VARIANCE = 0.2635

PROBABILITY OF HAVING LESS THAN      5.60 PPM = 0.1098
PROBABILITY OF HAVING LESS THAN      5.50 PPM = 0.0784

TIME OF TRAVEL IN DAYS = 4

PROBABILITY MASS DENSITIES FOR POLLUTION
PROBABILITY OF HAVING     0.80 PPM = 0.0000
PROBABILITY OF HAVING     0.90 PPM = 0.0001
PROBABILITY OF HAVING     1.00 PPM = 0.0002
PROBABILITY OF HAVING     1.10 PPM = 0.0004
PROBABILITY OF HAVING     1.20 PPM = 0.0009
PROBABILITY OF HAVING     1.30 PPM = 0.0019
PROBABILITY OF HAVING     1.40 PPM = 0.0036
PROBABILITY OF HAVING     1.50 PPM = 0.0064
PROBABILITY OF HAVING     1.60 PPM = 0.0106
PROBABILITY OF HAVING     1.70 PPM = 0.0164
PROBABILITY OF HAVING     1.80 PPM = 0.0241
PROBABILITY OF HAVING     1.90 PPM = 0.0304
PROBABILITY OF HAVING     2.00 PPM = 0.0438
PROBABILITY OF HAVING     2.10 PPM = 0.0546
PROBABILITY OF HAVING     2.20 PPM = 0.0647
PROBABILITY OF HAVING     2.30 PPM = 0.0731
PROBABILITY OF HAVING     2.40 PPM = 0.0789
PROBABILITY OF HAVING     2.50 PPM = 0.0816
PROBABILITY OF HAVING     2.60 PPM = 0.0808
PROBABILITY OF HAVING     2.70 PPM = 0.0769
PROBABILITY OF HAVING     2.80 PPM = 0.0703
PROBABILITY OF HAVING     2.90 PPM = 0.0619
PROBABILITY OF HAVING     3.00 PPM = 0.0525
PROBABILITY OF HAVING     3.10 PPM = 0.0430
PROBABILITY OF HAVING     3.20 PPM = 0.0340
PROBABILITY OF HAVING     3.30 PPM = 0.0260
```

Table 4.3.1 Continued

```
PROBABILITY OF HAVING    3.40 PPM = 0.0182
PROBABILITY OF HAVING    3.50 PPM = 0.0138
PROBABILITY OF HAVING    3.60 PPM = 0.0096
PROBABILITY OF HAVING    3.70 PPM = 0.0064
PROBABILITY OF HAVING    3.80 PPM = 0.0042
PROBABILITY OF HAVING    3.90 PPM = 0.0027
PROBABILITY OF HAVING    4.00 PPM = 0.0017
PROBABILITY OF HAVING    4.10 PPM = 0.0010
PROBABILITY OF HAVING    4.20 PPM = 0.0006
PROBABILITY OF HAVING    4.30 PPM = 0.0003
PROBABILITY OF HAVING    4.40 PPM = 0.0002
PROBABILITY OF HAVING    4.50 PPM = 0.0001
PROBABILITY OF HAVING    4.60 PPM = 0.0001
PROBABILITY OF HAVING    4.70 PPM = 0.0000

MEAN = 2.57     VARIANCE = 0.2387

PROBABILITY OF HAVING MORE THAN    3.20 PPM = 0.0859
PROBABILITY OF HAVING MORE THAN    3.10 PPM = 0.1199



TIME OF TRAVEL IN DAYS = 4

PROBABILITY OF HAVING    8.40 PPM = 0.0000
PROBABILITY OF HAVING    8.30 PPM = 0.0001
PROBABILITY OF HAVING    8.20 PPM = 0.0003
PROBABILITY OF HAVING    8.10 PPM = 0.0009
PROBABILITY OF HAVING    8.00 PPM = 0.0020
PROBABILITY OF HAVING    7.90 PPM = 0.0041
PROBABILITY OF HAVING    7.80 PPM = 0.0077
PROBABILITY OF HAVING    7.70 PPM = 0.0132
PROBABILITY OF HAVING    7.60 PPM = 0.0211
PROBABILITY OF HAVING    7.50 PPM = 0.0313
PROBABILITY OF HAVING    7.40 PPM = 0.0434
PROBABILITY OF HAVING    7.30 PPM = 0.0563
PROBABILITY OF HAVING    7.20 PPM = 0.0688
PROBABILITY OF HAVING    7.10 PPM = 0.0793
PROBABILITY OF HAVING    7.00 PPM = 0.0865
PROBABILITY OF HAVING    6.90 PPM = 0.0895
PROBABILITY OF HAVING    6.80 PPM = 0.0880
PROBABILITY OF HAVING    6.70 PPM = 0.0825
PROBABILITY OF HAVING    6.60 PPM = 0.0738
PROBABILITY OF HAVING    6.50 PPM = 0.0631
PROBABILITY OF HAVING    6.40 PPM = 0.0517
PROBABILITY OF HAVING    6.30 PPM = 0.0406
PROBABILITY OF HAVING    6.20 PPM = 0.0306
PROBABILITY OF HAVING    6.10 PPM = 0.0222
PROBABILITY OF HAVING    6.00 PPM = 0.0155
PROBABILITY OF HAVING    5.90 PPM = 0.0105
PROBABILITY OF HAVING    5.80 PPM = 0.0068
PROBABILITY OF HAVING    5.70 PPM = 0.0043
PROBABILITY OF HAVING    5.60 PPM = 0.0026
PROBABILITY OF HAVING    5.50 PPM = 0.0015
PROBABILITY OF HAVING    5.40 PPM = 0.0009
PROBABILITY OF HAVING    5.30 PPM = 0.0005
PROBABILITY OF HAVING    5.20 PPM = 0.0003
PROBABILITY OF HAVING    5.10 PPM = 0.0001
PROBABILITY OF HAVING    5.00 PPM = 0.0001
PROBABILITY OF HAVING    4.90 PPM = 0.0000
```

Table 4.3.1 Continued

```
MEAN = 6.84      VARIANCE = 0.2062

PROBABILITY OF HAVING LESS THAN  6.40 PPM = 0.1365
PROBABILITY OF HAVING LESS THAN  6.30 PPM = 0.0959


TIME OF TRAVEL IN DAYS = 5

PROBABILITY MASS DENSITIES FOR POLLUTION
PROBABILITY OF HAVING    0.40 PPM = 0.0000
PROBABILITY OF HAVING    0.50 PPM = 0.0001
PROBABILITY OF HAVING    0.60 PPM = 0.0004
PROBABILITY OF HAVING    0.70 PPM = 0.0011
PROBABILITY OF HAVING    0.80 PPM = 0.0026
PROBABILITY OF HAVING    0.90 PPM = 0.0055
PROBABILITY OF HAVING    1.00 PPM = 0.0104
PROBABILITY OF HAVING    1.10 PPM = 0.0179
PROBABILITY OF HAVING    1.20 PPM = 0.0282
PROBABILITY OF HAVING    1.30 PPM = 0.0411
PROBABILITY OF HAVING    1.40 PPM = 0.0553
PROBABILITY OF HAVING    1.50 PPM = 0.0695
PROBABILITY OF HAVING    1.60 PPM = 0.0817
PROBABILITY OF HAVING    1.70 PPM = 0.0902
PROBABILITY OF HAVING    1.80 PPM = 0.0939
PROBABILITY OF HAVING    1.90 PPM = 0.0924
PROBABILITY OF HAVING    2.00 PPM = 0.0863
PROBABILITY OF HAVING    2.10 PPM = 0.0766
PROBABILITY OF HAVING    2.20 PPM = 0.0647
PROBABILITY OF HAVING    2.30 PPM = 0.0522
PROBABILITY OF HAVING    2.40 PPM = 0.0403
PROBABILITY OF HAVING    2.50 PPM = 0.0299
PROBABILITY OF HAVING    2.60 PPM = 0.0212
PROBABILITY OF HAVING    2.70 PPM = 0.0145
PROBABILITY OF HAVING    2.80 PPM = 0.0095
PROBABILITY OF HAVING    2.90 PPM = 0.0060
PROBABILITY OF HAVING    3.00 PPM = 0.0037
PROBABILITY OF HAVING    3.10 PPM = 0.0022
PROBABILITY OF HAVING    3.20 PPM = 0.0012
PROBABILITY OF HAVING    3.30 PPM = 0.0007
PROBABILITY OF HAVING    3.40 PPM = 0.0004
PROBABILITY OF HAVING    3.50 PPM = 0.0002
PROBABILITY OF HAVING    3.60 PPM = 0.0001
PROBABILITY OF HAVING    3.70 PPM = 0.0000

MEAN = 1.87     VARIANCE = 0.1807

PROBABILITY OF HAVING MORE THAN  2.40 PPM = 0.0896
PROBABILITY OF HAVING MORE THAN  2.30 PPM = 0.1300


TIME OF TRAVEL IN DAYS = 5

PROBABILITY MASS DENSITIES FOR OXYGEN
PROBABILITY OF HAVING    8.70 PPM = 0.0001
PROBABILITY OF HAVING    8.60 PPM = 0.0002
PROBABILITY OF HAVING    8.50 PPM = 0.0007
PROBABILITY OF HAVING    8.40 PPM = 0.0020
PROBABILITY OF HAVING    8.30 PPM = 0.0048
PROBABILITY OF HAVING    8.20 PPM = 0.0100
```

Table 4.3.1 Continued

```
PROBABILITY OF HAVING    8.10 PPM = 0.0184
PROBABILITY OF HAVING    8.00 PPM = 0.0305
PROBABILITY OF HAVING    7.90 PPM = 0.0457
PROBABILITY OF HAVING    7.80 PPM = 0.0626
PROBABILITY OF HAVING    7.70 PPM = 0.0789
PROBABILITY OF HAVING    7.60 PPM = 0.0921
PROBABILITY OF HAVING    7.50 PPM = 0.1000
PROBABILITY OF HAVING    7.40 PPM = 0.1015
PROBABILITY OF HAVING    7.30 PPM = 0.0966
PROBABILITY OF HAVING    7.20 PPM = 0.0867
PROBABILITY OF HAVING    7.10 PPM = 0.0734
PROBABILITY OF HAVING    7.00 PPM = 0.0589
PROBABILITY OF HAVING    6.90 PPM = 0.0449
PROBABILITY OF HAVING    6.80 PPM = 0.0325
PROBABILITY OF HAVING    6.70 PPM = 0.0225
PROBABILITY OF HAVING    6.60 PPM = 0.0149
PROBABILITY OF HAVING    6.50 PPM = 0.0094
PROBABILITY OF HAVING    6.40 PPM = 0.0057
PROBABILITY OF HAVING    6.30 PPM = 0.0033
PROBABILITY OF HAVING    6.20 PPM = 0.0019
PROBABILITY OF HAVING    6.10 PPM = 0.0010
PROBABILITY OF HAVING    6.00 PPM = 0.0005
PROBABILITY OF HAVING    5.90 PPM = 0.0003
PROBABILITY OF HAVING    5.80 PPM = 0.0001
PROBABILITY OF HAVING    5.70 PPM = 0.0001
PROBABILITY OF HAVING    5.60 PPM = 0.0000

MEAN = 7.38     VARIANCE = 0.1575

PROBABILITY OF HAVING LESS THAN  7.00 PPM = 0.1370
PROBABILITY OF HAVING LESS THAN  6.90 PPM = 0.0921
```

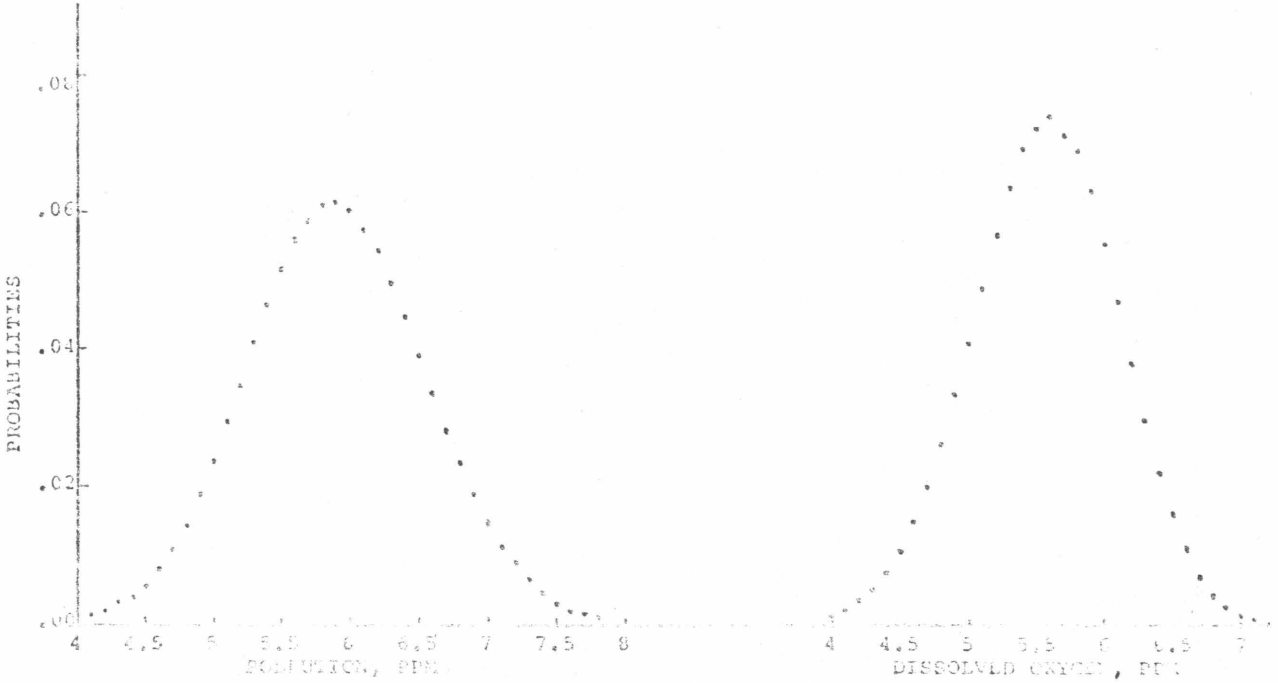

Figure 4.3.1. Pollution and Dissolved Oxygen
Mass Densities for t = 2

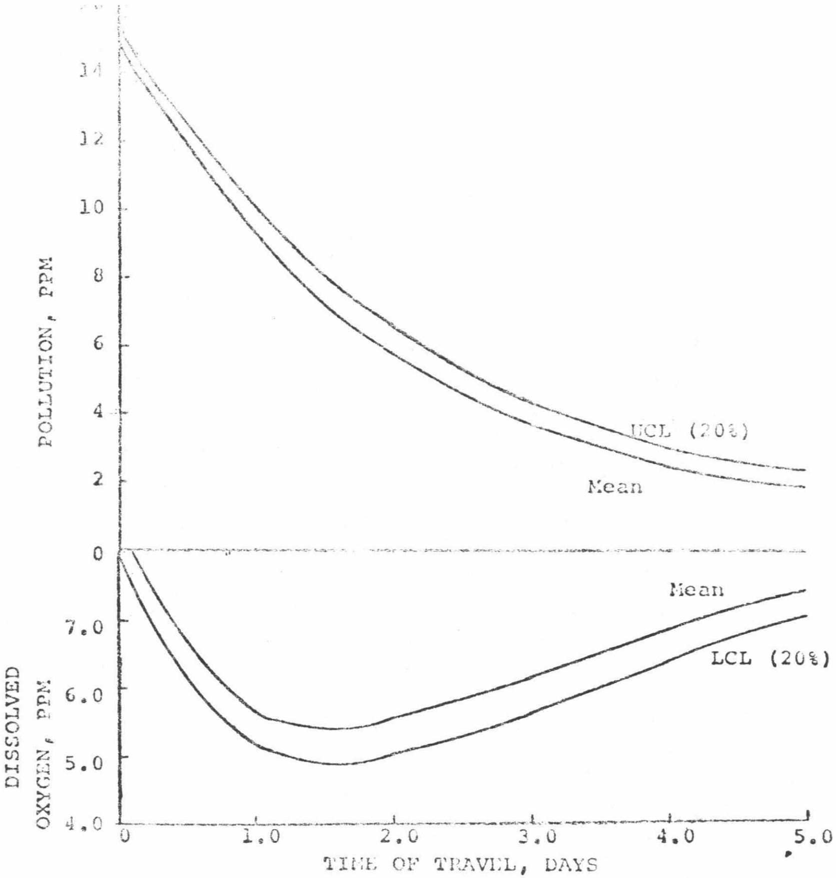

Figure [illegible].3.2. Pollution and Dissolved Oxygen Means and Confidence Limits

For pollution the upper 10 per cent confidence limit is plotted. It indicates that the variability is smallest at the extremes and larger in the middle. The problem is that in actual practice the variability in the measuring technique at low BOD levels is so large that it "masks" the real variability. This is not a serious problem because the prediction of pollution is much less important than the prediction of the dissolved oxygen.

For the dissolved oxygen the lower confidence limit is plotted. Fortunately the measurement of oxygen level is done with a small error in comparison with the actual variability in the river. Thus the measurement error is assumed to be zero. Due to its size, the river has a relatively small oxygen variance. In a smaller river where there is more oxygen variability the effect of measurement error might be less.

These results give a complete picture of what can be expected in the future when the initial conditions are realized. After the analysis, those in authority can make intelligent decisions on the course of action to be taken, i.e., decisions regarding waste-treatment plants, pollution laws, etcetera.

Consider again the discussion in Section 2.1 with regard to equation (2.1.4). In that equation the term containing $P_{M,O}(t)$ was said to be negligible in practical situations. If $P_{M,O}(t)$ is summed over M, one obtains the marginal probability for dissolved oxygen. The computer printout in this section shows that the probabilities below 3 ppm or state 30 are very small. Therefore the marginal probability that the oxygen is in state zero is negligible. Hence each term in the joint probability must be small, that is, $P_{M,O}(t) \doteq 0$.

V. LABORATORY TESTS

5.1 Apparatus

In order to further test the theoretical results, it was decided to run a series of controlled experiments in the laboratory. There are some advantages in a test of this kind. Most important is the control that one obtains. All sources of pollution are known. All changes in the amount of water are known. The temperature is kept constant at 21°C so that corrections of K_1 and K_2 with temperature changes are not needed. The amount of turbulence and hence K_2 is kept constant, whereas in a river there are rapids and there are quiescent conditions. Also, there is no time lapse while samples are being transported. Finally, the parameters are more easily measured as will be shown later.

On the other hand there is always the question of the applicability of laboratory results to actual conditions. It is difficult to simulate all that is occurring in the stream. For example, there was no benthal layer and hence no benthal demand. Although there was no sedimentation of pollution, the factor K_3 was not zero because the pollution was consumed in ways other than by bacterial action, K_1. Reasons for this will be given later.

The experiment was carried out in polyethylene waste cans which contained 20 liters of liquid. Each can represents a parcel of river water to which pollution is added. In a river situation this parcel would flow down stream, during which time the bacteria would consume the pollution and the loss of oxygen would be replaced by reaeration from the atmosphere. Of course in a river any parcel of water will mix with the surrounding water, but the surrounding water has essentially the same amount of pollution and dissolved oxygen as the parcel. Thus, after any time of travel one should obtain from any sample the same pollution and oxygen measurements as would be in the original parcel after the same time of travel, if one could recover it. The waste cans represented a parcel of water, the passage of time corresponding to movement downstream.

For this experiment five identical cans were used. The depth of liquid was 12 inches with an exposed surface area of about 90 square inches. The water in a river is mixed due to rapids, boats, currents, etcetera. The water of higher dissolved oxygen concentration near the surface is mixed with the underlying layers. The water in the cans also had to be agitated in some manner. The reaeration parameter K_2 depends on the agitation so that keeping the agitation constant keeps K_2 constant. This, of course, can only be approximated in an actual river situation. It was decided therefore to provide each can with a mechanical stirrer. A small motor was mounted above each can and a paddle extended into the can to a depth of 9 1/2 inches. The motor was adjusted horizontally so that the paddle pierced the water surface at the intersection of the diagonals of the approximately rectangular water surface. With the motors in operation the paddle stem was in the center of the vortex.

At the start of a run each can contained initially 20 liters of water from the pond on the campus of the Virginia Polytechnic Institute. This was done because it was sure to contain a complex bacterial flora and certain types would multiply rapidly when the substrate was added. Tap water could not be used because the chlorine would hinder the growth of the bacteria. The substrate chosen was the simple sugar, dextrose. To provide a complete diet for the bacteria this was supplemented with ammonium sulfate, the amount of which was such that there would be one part nitrogen to forty parts sugar. Also, potassium phosphate was added to provide one part phosphorous to 200 parts sugar.

It is known that the bacteria need these two elements in order to develop new cells, that is, they are constituents of cell protoplasm just as are carbon, hydrogen, and oxygen. The carbon is available in the sugar and the hydrogen and oxygen are present in the sugar and in water.

The results reported in this chapter are from two complete runs. Fresh pond water was used for each run.

5.2 Estimation of Parameters

During the course of a run periodic estimates were made of the pollution remaining and the amount of dissolved oxygen present. A discussion of all of the methods employed is included in Appendix III.

Throughout this dissertation the pollution has been defined as the Biochemical Oxygen Demand. With a relatively pure substrate such as the sugar the correlation between the BOD and another measurement the COD has been found to be high. The COD or Chemical Oxygen Demand has the advantage that the entire determination can be made in five hours whereas the BOD requires a five-day storage period before the determination is completed. A comparison of BOD and COD measurements at several levels of pollution showed the two methods give very similar results. All pollution measurements reported will be COD measurements.

In later sections the data from two separate runs are reported. These give the actual observed means and variances of both pollution and dissolved oxygen. The predicted means and variances are given by the theory previously developed in this dissertation. For these predicted means and variances, estimates of the parameters are needed. The estimates will be chosen from the range of values indicated by the measurements. The parameters are not expected to vary much with time. The pond water may change slightly, but its main purpose is to provide a source of bacteria which will grow and consume the sugar. Pollution is provided mainly by the sugar and not the pond water. The initial condition parameters which are determined by the experimenter are different for the two runs.

The first parameter to be considered is K_1+K_3. This is the rate at which the pollution decreases. The first question to be answered concerns the necessity of K_3. Why don't the pollution and oxygen decrease at the same rate K_1? This is because the bacteria use the sugar in two different ways. In the first case the bacteria take the sugar and completely oxidize it to carbon dioxide and water. This process of respiration uses most of the oxygen and provides energy for the cell. In the second process the bacteria use the energy to synthesize the sugar into cellular material; the cell grows and then splits into two cells which continue the process. Growths were evident on the sides of the can during the run. Thus all the pollution was not completely oxidized and hence both K_1 and K_3 are needed.

To estimate K_1+K_3, two 10 liter glass jars were filled with pond water and sugar was added. The jars were stirred with a magnetic stirrer. Several days elapsed while the bacteria developed. At the start of a time period the pollution was measured. The parameter, K_1+K_3, is estimated using formula (2.2.8)

$$L(t) = L_A e^{-(K_1+K_3)t} .$$

Solving for the parameter, one obtains

$$(K_1+K_3) = -\frac{1}{t}\ln(L_A/L(t)).$$

The results of the measurements are in Table 5.2.1.

Table 5.2.1 Data for Estimating the Parameter K_1+K_3

	Jar 1	Jar 2
Time, days	1.71	1.71
L_A, ppm	43.0	55.66
$L(t)$, ppm	28.0	32.70
K_1+K_3, per day	.25	.31

We will use the value .26 as our estimate of K_1+K_3. It is quite reasonable and gives the best fit.

The second parameter K_1 is the rate at which the oxygen is used by the bacteria. The jars of the previous section were used to estimate this parameter also. The procedure is to aerate the bottle by bubbling air through it so that there is plenty of dissolved oxygen at the start of the test. Then the initial dissolved oxygen concentration is measured. The bottle is then sealed with a stopper and water seal so that no air can enter. After a period of time the dissolved oxygen concentration will be reduced by the bacteria. The parameter is estimated using formula (2.2.14) with $K_2=0$. (In setting $K_2=0$, it is necessary to recall that

$$\lim_{K_2\to 0} \frac{1-e^{-K_2 t}}{K_2} = 0.$$

This can be shown by l'Hospital's Rule.)

$$C(t) = C_0 - \frac{L_A K_1}{K_1+K_3}\,(1-e^{-(K_1+K_3)t}).$$

Solving for K_1 and replacing K_1+K_3 by its estimate .26, one obtains

$$K_1 = [\frac{C_0-C(t)\quad(.26)}{L_A(1-e^{-.26t})}].$$

Several different trials yielded the results of Table 5.2.2.

Table 5.2.2 Estimates of the Parameter K_1

t, days	C_o, ppm	C(t), ppm	L_A, ppm	K_1, per day
8.5/24	8.00	7.10	35	.076
13/24	7.55	5.85	34	.099
10/24	7.70	4.70	44.7	.167
10/24	7.80	4.80	54.5	.137

It was found in the later sections that K_1=.085 fit the first set of data and K_1=.07 fit the second set of data. Both values are reasonably close to the above estimates. The two values are close together and the difference could be due to any one of many variables.

The final rate parameter to be estimated is the reaeration rate K_2. In order to do this the water was boiled in order to kill the bacteria (so that K_1 and K_1+K_3 would be zero) and to lower the dissolved oxygen concentration. Then the water was allowed to cool overnight in "stoppered" vessels so that no reaeration could take place. Finally the water was put in the waste cans again and the mechanical stirrer was started. An initial dissolved oxygen determination was made.

There are several ways in which one can get misleading data from this experiment. First of all, it is impossible to get rid of all of the bacteria, so that while the water is cooling and reaerating, if there is sugar present the bacteria will again develop and the K_1 and K_1+K_3 will no longer be zero. Also K_2 must be measured using dissolved oxygen concentrations approximating those to be found in the cans. For example the reaeration rate for concentrations between 3 and 5 ppm is slightly greater than the reaeration rate obtained for concentrations between 5 and 7 ppm. However the formulas are applied only in regions above 5 ppm, since 3 ppm is not considered enough dissolved oxygen for optimum use of most natural waters.

To overcome these difficulties, fresh pond water (with no sugar present) was boiled and cooled. It was then placed in the can and aerated rapidly until it reached around 5 ppm. Then some pollution was added, the stirrer was started, and the concentration of dissolved oxygen was determined. In this situation there are relatively few bacteria present and it takes time for them to get acclimated to the environment and develop. The reaeration process is affected by the presence of solutes. Thus in order to evaluate reaeration correctly it must be carried out in the presence of the sugar substrate.

For K_1 and K_1+K_3 zero, formula (2.2.14) for dissolved oxygen with fixed initial condition becomes

$$N(t) = C_N' - (C_N' - C_o)e^{-K_2 t} .$$

After rearranging and taking logarithms one obtains

$$\ln (C_N' - N(t)) = \ln (C_N' - C_o) - K_2 t.$$

The maximum dissolved oxygen concentration $C_N' = 8.4$ (this will be justified subsequently.) Using four pair of data points $(t_i, N(t_i))$ one can estimate using simple linear regression K_2 and $\ln (C_N' - C_o)$ and therefore C_o. More than four points are desirable but the volume has to remain constant.

Table 5.2.3 Data for Estimating the Parameter K_2

t_i, days	$N(t_i)$,ppm
0	6.3
9/24	7.4
18/24	7.8
22/24	7.88

The estimates are $K_2 = 1.53$ and $C_o = 6.4$. In the first equation of this paragraph the error is assumed to be multiplicative. On taking logarithms the error becomes additive.

Using additional theory of simple linear regression one can obtain one standard deviation interval for K_2. For this case the σ limits are

$$1.38 < K_2 < 1.68.$$

The value chosen for K_2 is 1.6.

The correction for the maximum amount of dissolved oxygen will now be discussed. The maximum concentration that will dissolve in water at 21°C is 9 ppm at sea level. The correction is due to the elevation of Blacksburg above sea level.

It is generally assumed that the ratio of the solubilities at different barometric pressures equals the ratio of the barometric pressures, so that

$$\frac{C_{N'}}{9} = \frac{B}{760}$$

where B is the barometric pressure at Blacksburg in units of millimeters of Mercury and 760 is the barometric pressure at sea level (see Fair and Geyer (6)). Now B can be obtained from the formula

$$z = (17,350)\frac{760 - B}{760 + B}$$

which is in a reduced form because the temperature is constant. The variable z equals the difference in elevation in meters.

The elevation at Blacksburg is approximately 2200 feet which is equivalent to 660 meters. On performing the above algebraic manipulations one obtains $C_N' = 8.4$.

All of the parameters except Δ have been specified and the estimate of Δ can be easily obtained because the variance function is linear in Δ. The parameter Δ can be estimated from the ratio of the sample variance to the coefficient of Δ in the predicted variance. For the two sets of data that follow, we choose $\Delta = 0.0055$.

5.3 First Trial Run

Initially the five cans were filled with 20 liters of pond water. To each can 800 mg of sugar dextrose were added for an initial concentration of 40 ppm. For a better diet, ammonium sulfate was added in the ratio of 1 part nitrogen to 40 parts sugar, that is 800 x 1/40 x 4.72 = 94.4 mg, where the factor 4.72 is the ratio of the molecular weight of ammonium sulfate to the molecular weight of nitrogen present in the compound. Similarly, potassium phosphate was added in the ratio of 1 part phosphorous to 200 parts sugar yielding 800 x 1/200 x 6.85 = 27.4 mg, where the factor 6.85 is the reciprocal of the percentage of phosphorous in the molecular weight of the compound.

During the course of a run the DO and COD were measured periodically. A description of the tests involved is in Appendix III.

The COD readings must be corrected for the chloride in the pond water. The chloride increases the COD but the bacteria do not oxidize it, so it is necessary to subtract its effect on the COD determination. To determine the amount present the Mohr Method was used. This method is described in Appendix III. The average of two measurements yielded a concentration of 21 mg/l. The oxygen demand equivalent to this is found by multiplying the mg/l Cl by 0.23. This gives a value of 4.83 and this value was subtracted from all COD values reported in this run.

Before the parameters can be used to give the predicted means, it is necessary to establish a point, which will be the reference point $t = 0$. In a river situation, the pollution comes in almost continuously, from a sewage treatment plant for example. There are then in the river some distance below the outfall, an abundance of bacteria. These types will readily attack the effluent. The pollution is used up at the rate $K_1 + K_3$ not long after it enters the river. This is somewhat similar to the case with the trash cans. In the waste cans, time is required while the bacteria become acclimated to the new environment. Then the particular strains of bacteria that attack sugar must develop. There is a period of time during which K_1 and $K_1 + K_3$ are less than their ultimate values. The decrease in pollution should theoretically be largest at the beginning because the rate is proportional to the amount present. Now, consider Figure 5.3.2, which is a plot of pollution for various times of travel. The decrease from Monday night to Tuesday night is small compared with the decrease from Tuesday night to Wednesday night. If $K_1 + K_3$ were at its ultimate value, the latter decrease should be less than the former. Thus, $t=0$ will be at 9 PM Tuesday.

The observed DO values plotted against time of travel are shown in Figure 5.3.1. The time $t = 0$ has been established as 9 PM Tuesday. The average of the five points at $t = 0$ in Figure 5.3.1 give the initial concentration of dissolved oxygen C_o as 6.4. Similarly, in Figure 5.3.2 the average at $t = 0$ gives $L_A = 52$ ppm.

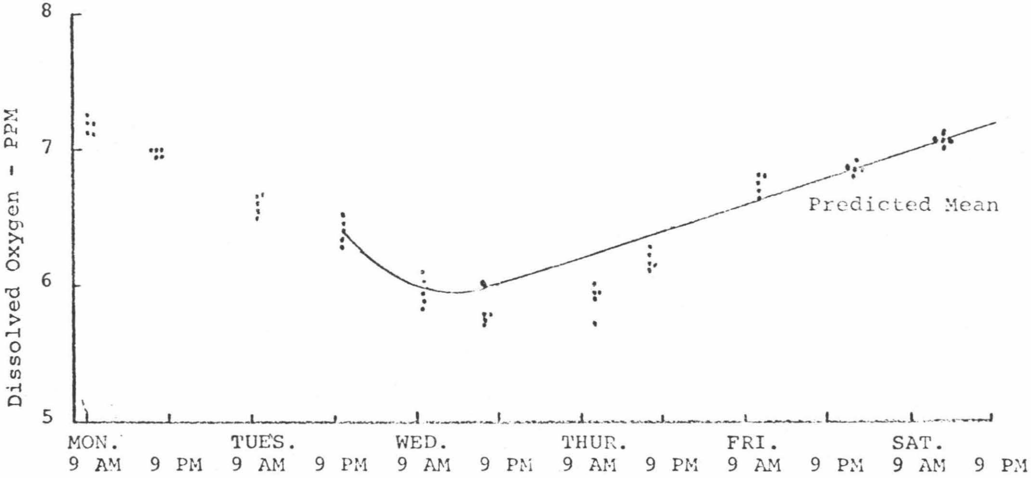

Figure 5.3.1. Predicted and Observed Dissolved Oxygen-Run 1

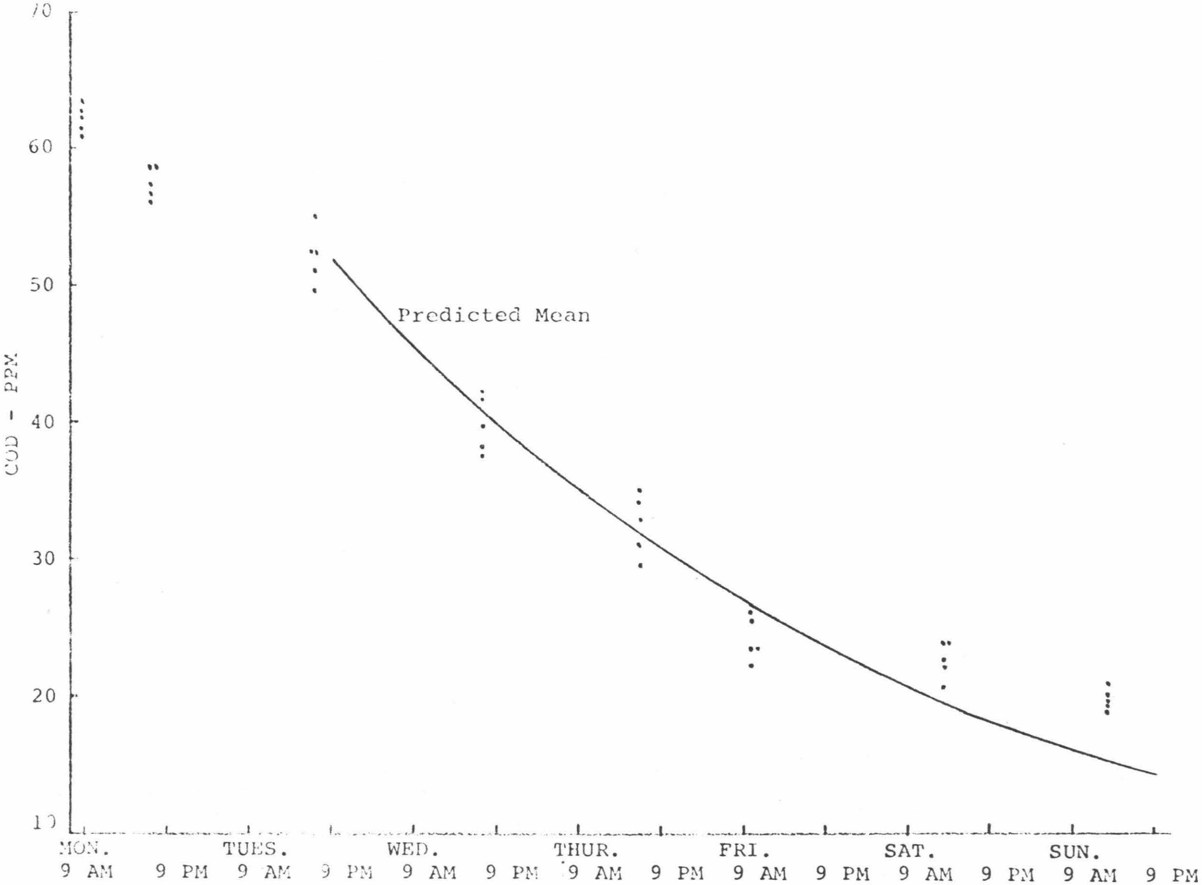

Figure 5.3.2. Predicted and Observed COD - Run 1

The predicted mean for dissolved oxygen is given in formula (2.2.14)

$$C(t) = C_N' - (C_N'-C_o)e^{-K_2t} - \frac{L_AK_1}{K_2-(K_1+K_3)}(e^{-(K_1+K_3)t}$$

$$-e^{-K_2t}).$$

This formula is in reduced form because L_a and D_B have been set equal to zero. The predicted pollution equation is formula (2.2.8)

$$L(t) = L_Ae^{-(K_1+K_3)t}.$$

In this formula L_a has been set equal to zero.

The values of the rate parameters K_1, K_2 and K_1+K_3 were selected from within the estimated ranges given in Section 5.2. Midpoints or averages were used initially and then these were adjusted slightly to give a better fit of the means. The "estimated" values of the parameters are

K_1+K_3 = 0.26 per day

K_2 = 1.6 per day

K_1 = 0.085 per day.

The predicted means are the continuous curves in Figures 5.3.1 and 5.3.2. The predicted curves fit the observed values very well.

The dissolved oxygen measurements were taken in the following manner. All five cans were tested with the DO meter. Then one of the cans (a different one each time) was tested using the Winkler Method and the scale factor ϕ was obtained. The Winkler Method is described in Appendix III. All meter readings were adjusted by this factor ϕ. Any slight error in the Winkler Method would raise or lower all five points. This might possibly account for the readings on Thursday at 9 AM being slightly low.

The sample variances for pollution are in Table 5.3.1.

Table 5.3.1 Sample Pollution Variances - Run 1

Time of Travel, Days	Sample Variance
0.00	3.92
0.92	4.13
1.88	5.01
2.54	2.67
3.71	1.91
4.71	0.61

As stated above the variances increase with time and then decrease. The observed and predicted maxima occur at approximately the same time of travel.

The next general discussion concerns the variances of both pollution and dissolved oxygen. Before considering specific values for the estimated and predicted variances, consider all of the points in Figure 5.3.1. and Figure 5.3.2. The theory developed in this dissertation predicts that for fixed initial conditions, such as were the case on Sunday evening, the variances of both pollution and dissolved oxygen increase to a maximum and then decrease with time. This is the impression that one gets looking at the dispersion of points. Formulas for the predicted variance are (2.2.9) and (2.2.15).

Unfortunately, the measurement error in the pollution is of such magnitude that it masks the variability between the cans. The standard error of measurement was found to be about 2.0. This value was obtained from a test using five replicates. Therefore, only the oxygen variances can be checked in any detail.

The theory predicts the pollution variability to be much larger than the oxygen variability, but in this case the measurement variability of the pollution is still too large to be ignored. In actual stream conditions one is not likely to run into this problem since the stream variability will be of much greater magnitude than one finds in the controlled experiment. Thus the measurement errors will be much less significant. In using these results one is most interested in the dissolved oxygen because it is the lack of dissolved oxygen which is harmful to the fish.

It has been previously established that $t = 0$ is at 9 PM on Tuesday. At this point it can no longer be said that the initial amounts of pollution and dissolved oxygen are fixed. Therefore, the theory of Chapter III, Section 3.2 and the computer program are needed. The formula for the dissolved oxygen variance is formula (3.2.9) of Chapter III. The parameters D_B and L_a are set equal to zero.

It is necessary to specify the mean level of pollution and dissolved oxygen and also the upper and lower limits of the range for each one. For the pollution a reasonable estimate of the limits would be $R_1 = 50$ ppm and $R_2 = 54$ ppm.

The estimates of the limits of the range of the dissolved oxygen are obtained in the following manner. The standard deviation at $t = 0$ is $s = 0.087$. The limits which are within 3 standard deviations of the mean are 6.1 ppm and 6.7 ppm for D_1 and D_2 respectively.

What is the measurement error associated with dissolved oxygen? The physical situation here obviously favors accurate readings. The meter is stationary throughout; the time lapse between readings is short, although several minutes are allotted each can for the meter to stabilize; the velocity of water passing the probe is constant. The meter is read to the nearest .05 ppm so that the maximum roundoff error is 0.00078 if each reading is off by 0.025. The readings are highly reproducible, rarely differing by .05. Scaling the roundoff error by the meter factor, the reciprocal of ϕ^2 one obtains about 0.001 as a reasonable order of magnitude for the variance. This is not large when compared with the variance between the cans.

Figure 5.3.3 gives the observed variances and the predicted dissolved oxygen variance curve. The observed values are reasonably close to the predicted curve and the sample points and predicted curve have maxima at approximately the same time of travel. The sample variances are based on only 5 observations.

Figure 5.3.4 is a control chart for the dissolved oxygen variances. The sample ranges appear to be in control. Some of the observed ranges are above the mean range and some below. Only the first observation is near the upper control limit.

Next it was decided to use all of the data in examining the variance. Before Tuesday at 9 PM the rates K_1 and $K_1 + K_3$ were not at their ultimate values. To compensate for this the time from Monday 9 AM to Tuesday 9 PM was shortened. It was found by trying several values that when t is replaced by -7/24 in the dissolved oxygen mean function, the value obtained is 7.2, and this value is the measured mean for Monday 9 AM. If t = 0 is taken at 2 PM Tuesday, the initial dissolved oxygen concentration is 7.2 and L_A, the initial pollution, must be changed to 56 ppm. The three rate parameters remain unchanged.

With t = 0 at 2 PM Tuesday consider the variance function for fixed initial conditions. The predicted variance is formula (2.2.15) of Chapter II. Figure 5.3.5 shows a curve of the predicted variance function with the sample variances plotted. It is evident that both have a tendency

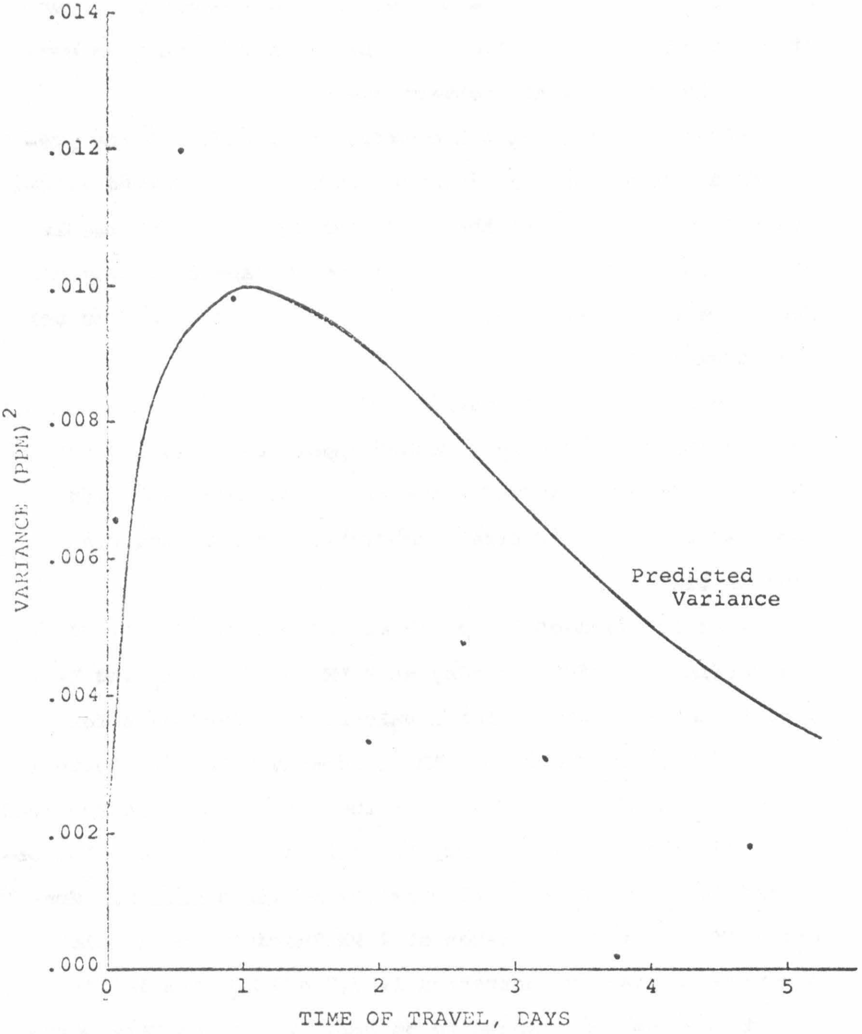

Figure 5.3.3. Predicted and Observed DO Variance - Run 1

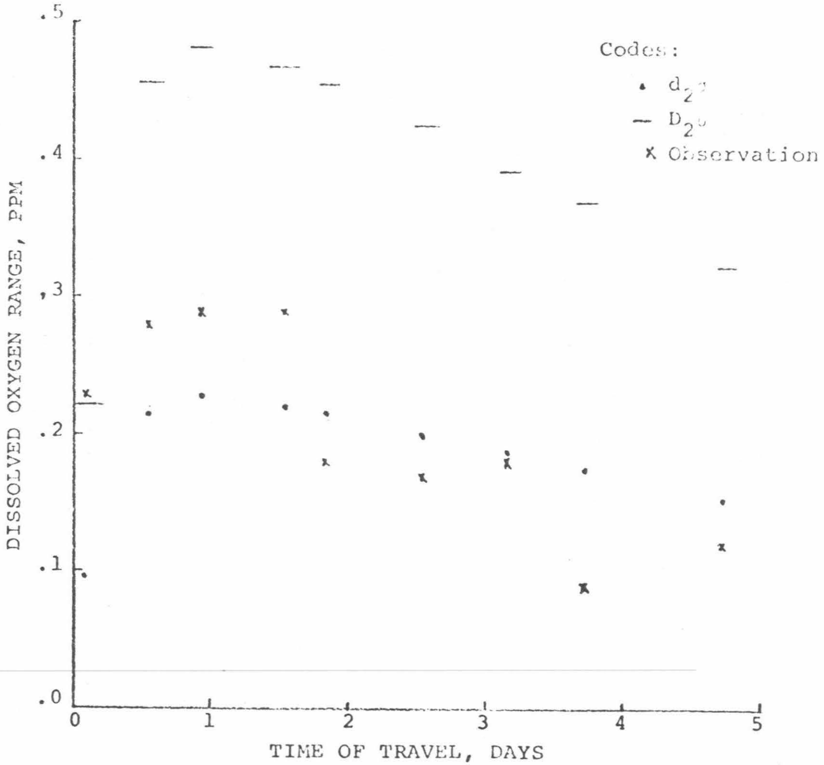

Figure 5.3.4. Control Chart for Ranges - Run 1

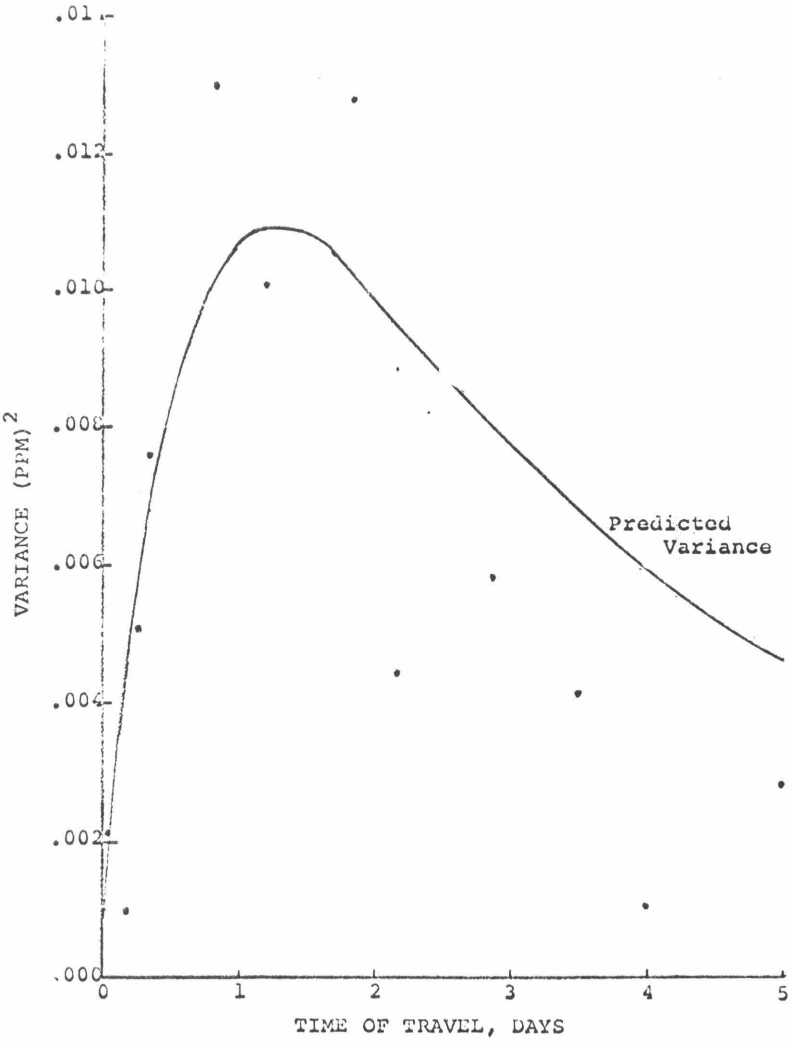

Figure 5.3.5. Predicted and Observed DO Variances - Run 1

to start at a small value, increase to a maximum and then decrease. The fit is good although there is some tendency for the observed points to have a higher maximum. The maxima for the observed and predicted variances seem to occur at approximately the same point in time.

5.4 Second Trial Run

The run to be discussed in this section is similar to the first run, but it is different in several significant ways. The first major difference is that in this run the pollution L_a coming in from the sides has been simulated. The method consisted of having a source of pollution, namely sugar dissolved in distilled water. This sugar solution was pumped through one channel of a feed pump to a "constant level" reservoir. Five tubes carried the water from the reservoir through five other channels of the feed pump and into the cans. The reservoir was kept at a constant level so that the pump always had the same amount of work to do and would always pump at the same rate. The flow rate in each channel was adjustable, so that the five channels were each calibrated to give the same flow. The flow chosen was 500 ml/day and this flow replaced losses due to sampling and evaporation. It was decided to choose L_a equal to 2 mg/l/day. This concentration corresponds to 40 mg in 20 liters. From the molecular weights of sugar and oxygen it was determined that 40 mg/20 l COD corresponded to 37.5 mg sugar in 20 l. Thus the pollution entering contained 37.5 mg sugar in 500 ml of distilled water. The pollution entering also contained ammonium sulfate and potassium phosphate in proportions as described earlier.

The second difference between this and the first run concerns the chloride ion which interferes with the COD test by oxidizing the chlorides to chlorine and thus using some of the potassium dichromate. In the first run its effect was subtracted from all COD measurements. A better method of handling the situation is to add mercuric sulfate which keeps the chloride in undissociated form, and this method was adopted for this run. To correct for a 100 mg/l concentration of chloride in a 50 ml sample the sample could contain 5 mg chloride as a maximum. It is then recommended that mercuric sulfate be added in a 10 to 1 ratio. Therefore 50 mg of mercuric sulfate are needed. Mercuric sulfate can be dissolved in distilled water made slightly acidic with sulfuric acid. The solution is of a concentration such that only .3 ml are needed for each sample in order to get 50 mg of mercuric sulfate.

A third major difference between the two runs concerns the way in which they were begun. The principle was to add a small amount of sugar when the cans are filled to encourage the growth of bacteria. Approximately 24 hours later the major portion of the pollution is added. The initial dose doesn't affect the dissolved oxygen level appreciably, if it is not too large and when the major dose is added the variance of both pollution and dissolved oxygen are small. If the first dose is too small then the bacteria won't develop in a day. It was arbitrarily decided to make the initial small dose one tenth of the dose given in run 1, namely, 80 mg dextrose, 9.44 mg ammonium sulfate and 2.74 mg of potassium phosphate.

The second run began on Wednesday at 8 PM with the filling of the five cans and the starting of the five stirrers. The initial small dose of pollution mentioned in the preceding paragraph was added, but the pollution in the form of L_a was not started at this time. The cans were not touched for the next 24 hours while the bacteria developed.

At 8 PM on Thursday the main pollution dose was added to each can. The dose was identical to the pollution added for run 1, namely, 800 mg of dextrose, 94.4 mg of ammonium sulfate, and 27.4 mg of potassium phosphate for each can. The pump was started for the addition of pollution, L_a, and it continued in operation for the duration of the run.

During the course of the run, periodic measurements were made of the COD and DO. The methods involved in measuring the COD and DO are described in Appendix III.

Figure 5.4.1 is a plot of the COD against time of travel for the run. Before any prediction equations can be used, it is necessary to determine $t = 0$. From the data points it is clear that the decrease in pollution from Friday morning to Friday evening is small, so that K_1 is below its ultimate value. After Friday evening, it appears that K_1 has reached its ultimate value and therefore $t = 0$ was chosen as 8 PM Friday.

One can find the predicted COD mean from formula (2.2.8) in Chapter II. The value chosen for L_A is 61 ppm, the average of the five data points.

The data points for the dissolved oxygen profile for this run are plotted in Figure 5.4.2. The time for $t = 0$ has been established as 8 PM Friday. The predicted dissolved oxygen mean function is formula (2.2.14) in Chapter II. It is in reduced form because D_B has been set equal to zero. The initial dissolved oxygen concentration C_o is chosen at 7.6 ppm, the average of the five data points.

The estimates of the rate parameters used for the prediction equations are:

$$K_1 + K_3 = .26 \text{ per day}$$

$$K_2 = 1.6 \text{ per day}$$

$$K_1 = .07 \text{ per day.}$$

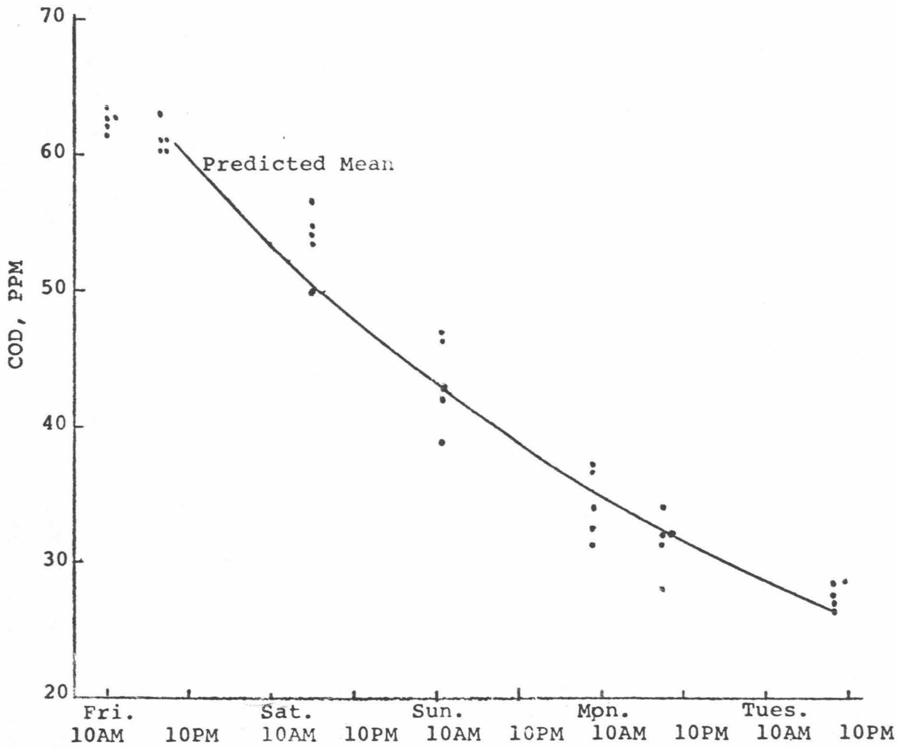

Figure 5.4.1. Predicted and Observed
COD Values - Run 2

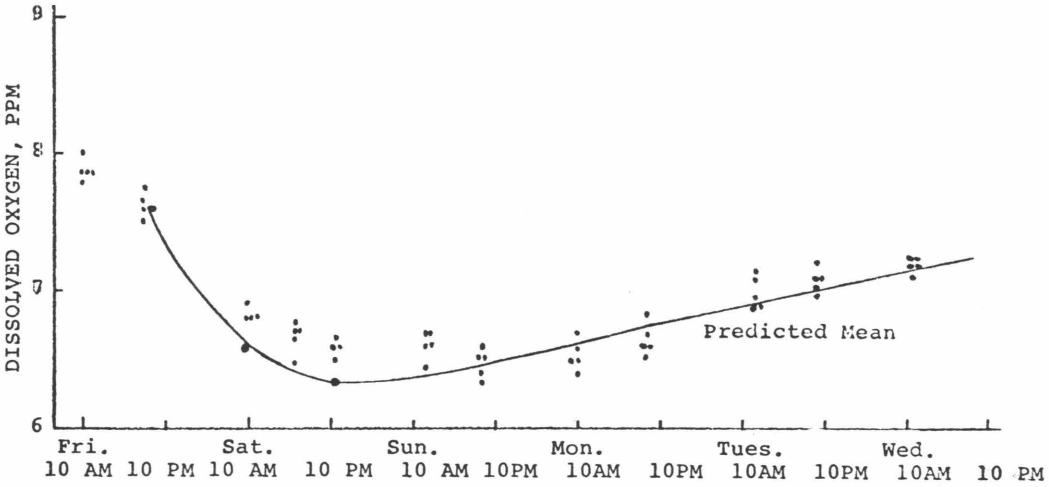

Figure 5.4.2. Predicted and Observed DO Values - Run 2

These are the same as the rate parameters of the previous section except that K_1 is slightly lower in this case.

The predicted mean functions, evaluated with the aid of the computer program, are plotted in Figures 5.4.1 and 5.4.2. The curves provide a very good fit of the observed data.

Next consider the variances for pollution and dissolved oxygen. With this run, as with the previous run, Figure 5.4.1 and Figure 5.4.2 indicate that the variances are larger near the center of the run than at either of the extremes. The form of the theoretical variance functions are described similarly. The magnitudes will now be examined.

The time t = 0 has already been established as 8 PM Friday. Since the initial variances are small the theory of Section 2.2 will be used and the predicted results will be obtained from the computer program of Section 2.3.

Again the size of the measurement error in the pollution measurements was a large portion of the total variation. Thus the theoretical and observed variances could not be compared.

The sample variances for the pollution measurements are in Table 5.4.1.

Table 5.4.1. Sample Pollution Variances - Run 2

Time of Travel, days	Sample Variance
0.00	1.22
0.83	6.06
1.63	10.67
2.54	6.58
3.00	4.77
4.00	2.80

With increasing time the sample variances rise to a maximum and then decrease. The predicted variance follows the same form. The dissolved oxygen variance formula is formula (2.2.15) of Chapter II. In the formula D_B has been set equal to zero.

The value of Δ used in this case is the same as the value used in the previous run, namely, 0.0055. Figure 5.4.3 shows the predicted variance curve and the observed variances. The sample estimates were obtained from samples of size five. The observed variances are larger near the middle of the time of travel scale and this is the same form as the continuous predicted curve.

In the previous run the observed variances had more of a tendency to "peak" than the predicted variances and in this run the opposite is true. This fact further substantiates the theoretical variance function.

Now let us examine the dissolved oxygen variances on a range control chart. Figure 5.4.4, the control chart, shows that the ranges are close to the expected values. None of the observed points are near the upper limits, since the probability of being above these limits is only 0.005.

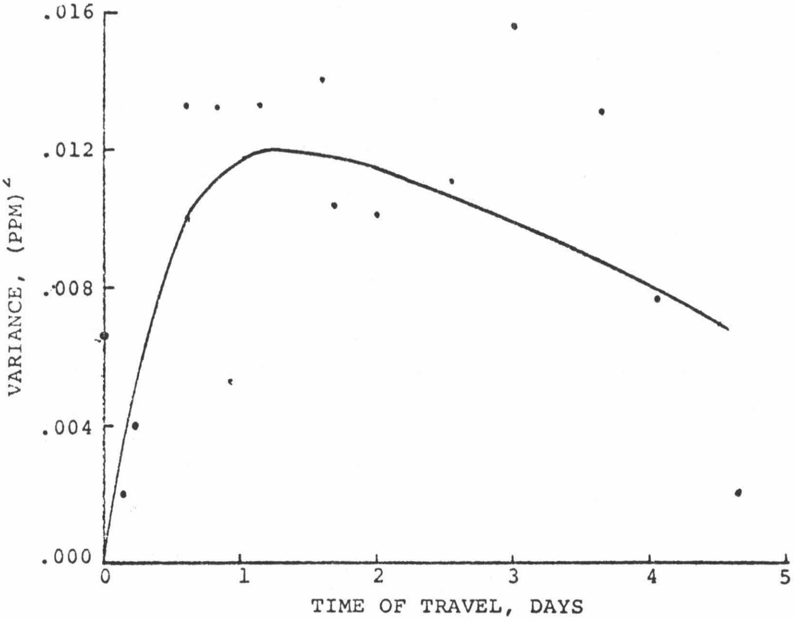

Figure 5.4.3. Predicted and Observed DO Variances - Run 2

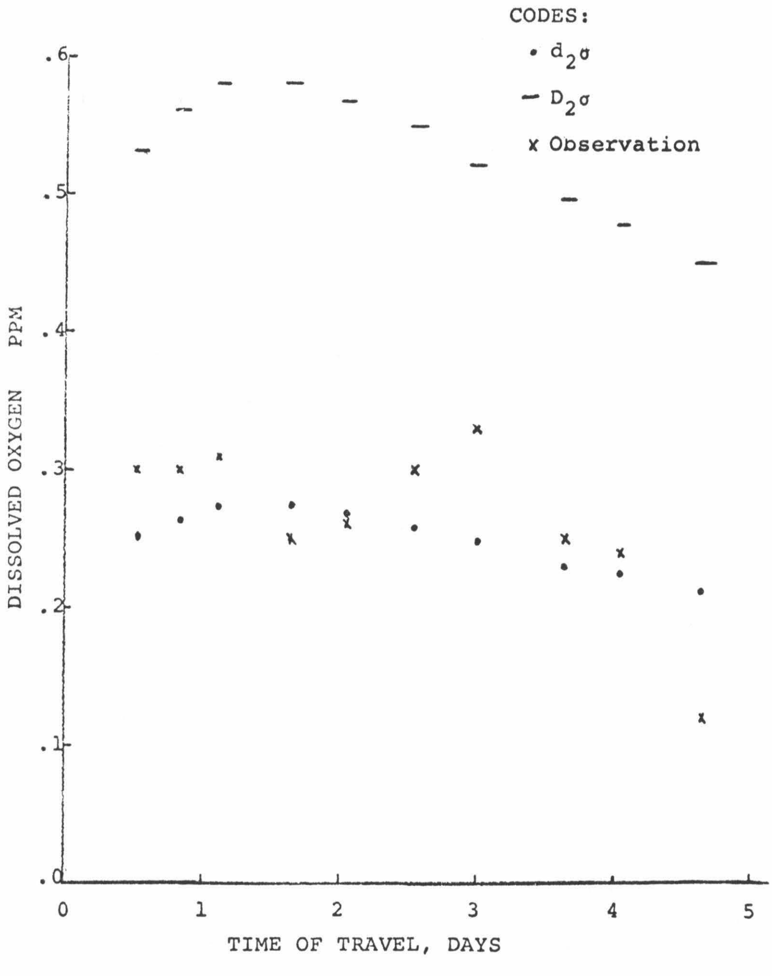

Figure 5.4.4. Control Chart for Ranges - Run 2

VI. CONCLUSIONS

6.1 Uses of the Results

In order to find out where these results can be applied, one need only refer back to Chapter I in which the uses of the river are partially enumerated. A municipality or industry can determine the effect of their pollution on the river. In designing a municipal or industrial waste treatment facility the designer can find out what the quality of his effluent should be, so as not to cause excess pollution.

The commercial fishing industry and sport fishermen want the fish protected and fish kills prevented. Government agencies can formulate minimum requirements in terms of the probability that the dissolved oxygen is below a fixed concentration a certain proportion of the time. For example, the requirement might be that ten percent of the time the dissolved oxygen can be below 5 ppm. Also the requirement might be stated as two probabilities associated with two concentrations instead of one probability and concentration as in the previous example.

Recreation agencies would want upper limits on the amount of pollution flowing through a park. The requirement could be stated in terms of the proportion of time that the pollution can exceed a given concentration.

One may be interested in varying the conditions. The initial conditions at some future time may be used to determine what the condition of the river would be at that time. Also changes in the rate parameters may be of interest. All of the above results could be obtained with the computer programs.

6.2 Conclusions

The problem investigated here has been considered by many people, but there were essentially no approaches to the problem from a stochastic point of view. With very few assumptions in addition to the rate assumptions assumed by others, the results were obtained. The only additional parameter to be estimated is Δ and its value is easily obtained from the variances of dissolved oxygen observations made during a stream survey. All of the arithmetic work involved can be accomplished with the computer programs provided.

The results obtained in this investigation agree with other proven results in so far as the proven results have been developed. The additional results of this investigation have been confirmed by the two sets of data examined. Also, the author of the Sacramento River study made an observation on the scatter of points along the dissolved oxygen profile. These results agreed with the theoretical results obtained here. Much data would be needed to test the form of the predicted probability mass densities, but it seems that a unimodal distribution is appropriate.

In regard to future work, the present theoretical results could be expanded if needed. Any set of initial conditions could be substituted into the procedure outlined in Section (2.1) and hence the probabilities, means and variances could be obtained. Higher moments could be obtained if desired. Also, an approximation to the probability distribution could be obtained which requires less calculation. Finally, but certainly not the least important, work should be done on the estimation of parameters. Some estimation procedures were briefly considered. These techniques require the solution of very difficult algebraic expressions.

Some work has been done and doubtless other work will follow in which the five assumptions in Section 2.1 will be changed or be increased in number. The methods used here will again be applicable. A derivation paralleling Section 2.1 will be needed and then results can be obtained for various initial conditions.

APPENDIX I COMPUTER PROGRAMS

PROGRAM 1

```
C       PROGRAM FOR FINDING PROB. DISTRIBUTIONS WITH FIXED
        INITIAL COND.  DIMENSION A(1000), B(1000), T(1000),
      1 UP(1000), D(1000), S(1000), E(1000)
        READ (5,7) ZK1, ZK2, ZK3, ZLA, ZDB, DELTA, ALPHA,I, NP, J
      7 FORMAT (7F6.0, 314)
        AI = I
        NJ = NP - J
        ANP = NP
        AJ = J
        ZK4 = ZK2 = (ZK1 + ZK3)
        DO 12 NT = 1, 5
        TIME = NT
        Z = (ZK1 + ZK3) * TIME
        U = EXP (-Z)
        V = 1.0 - U
        W = (ZLA * V) / (DELTA * (ZK1 + ZK3))
        X = EXP (-W)
        CALL BINOM (V, I, U, I+1, B, E)
        CALL EXPON (I, W, I+1, A)
        CALL SERMU (I, I, B, I+1, A, I+1, T, I+1)
        IP1 = I + 1
        DO 14 N = 1, IP1
        BNP = N - 1
        UP(N) = BNP * DELTA
     14 T(N) = X * T(N)
        SMEAN = (ZLA/ (ZK1 + ZK3)) * V + AI * DELTA * U
        SVAR = DELTA*(ZLA/(ZK1 + ZK3) + AI * DELTA * U) * V
        N = 1
        SUM1 = T(1)
      8 IF (1.0 - ALPHA - SUM1) 10, 10, 9
      9 N = N + 1
        SUM1 = SUM1 + T(N)
        GO TO 8
     10 SUM2 = SUM1 - T(N)
        SUM3 = 1.0 - SUM1
        SUM4 = 1.0 - SUM2
        WRITE (6,500) NT
    500 FORMAT (26HOTIME OF TRAVEL IN DAYS = I3/)
        WRITE (6,600)
    600 FORMAT (41H PROBABILITY MASS DENSITIES FOR POLLUTION)
        DO 60 KN = 1, IP1
        IF (T(KN) - 0.00001) 60, 60, 80
     80 WRITE (6,601) UP(KN), T(KN)
    601 FORMAT (23H PROBABILITY OF HAVING F7.2, 7H PPM = F7.4)
     60 CONTINUE
        WRITE (6,700) SMEAN, SVAR
    700 FORMAT (8HOMEAN = F7.2, 13H  VARIANCE = F9.4/)
        WRITE (6,800) UP(N), SUM3
    800 FORMAT (33H PROBABILITY OF HAVING MORE THAN F7.2,
      1  7H PPM = F7.4)
        WRITE (6,800) UP(N-1), SUM4
        SKT = (-ZK2)* TIME
        UPK = EXP(SKT)
        WPK = 1.0 - UPK
        F = (ZLA * ZK1)/(DELTA*(-ZK4))*(WPK/ZK2-V/(ZK1+ZK3))
       1 +ZDB * WPK / (DELTA * ZK2)
        G = ZK1 * (UPK - U) / (-ZK4)
```

```
      NP1 = NP + 1
      Y = 1.0 - G
      CALL BINOM (WPK, NJ, UPK, NJ + 1, B, E)
      CALL BINOM (Y, I, G, IP1, D, E)
      CALL EXPON (NP, F, NP1, A)
      CALL SERMU (NP, NJ, B, NJ + 1, A, NP1, S, NP1)
      CALL SERMU (NP, I, D, I +1, S, NP1 T, NP1)
      FE = EXP(-F)
      DO 26 N = 1, NP1
      BNP = NP1 - N
      UP(N) = BNP * DELTA
   26 T(N) = FE * T(N)
      SMEAN = DELTA * ANP - ZLA * ZK1/ ((ZK1 + ZK3) * ZK4) * V +
     1 ZLA * ZK1 / (ZK2 * ZK4) * WPK - ZDB / ZK2 * WPK - DELTA
     2 * (ANP - AJ) * UPK - AI * DELTA * ZK1 / ZK4*(U-UPK)
      SVAR = DELTA * (-(ZLA * ZK1)/ (ZK2 * ZK4) * WPK+ZLA*ZK1/
     ) ((ZK1 + ZK3) * ZK4) * V + ZDB / ZK2 * WPK + DELTA *
                                                      (ANP - AJ)
     2 * WPK * UPK + AI * DELTA * ZK1 / ZK4 * (U - UPK) *
                                                      (1.0 - ZK1/
     3 ZK4 * (U - UPK)))
      N = NP + 1
      SUM1 = T(N)
   18 IF (ALPHA - SUM1) 20, 20, 19
   19 N = N - 1
      SUM1 = SUM1 + T(N)
      GO TO 18
   20 SUM2 = SUM1 - T(N)
      WRITE (6,100) NT
  100 FORMAT (26HOTIME OF TRAVEL IN DAYS = I3/)
      WRITE (6,101)
  101 FORMAT (38H PROBABILITY MASS DENSITIES FOR OXYGEN)
      DO 70 JK = 1, NP1
      IF (T(JK) - 0.00001) 70, 70, 90
   90 WRITE (6,102) UP(JK), T(JK)
  102 FORMAT (23H PROBABILITY OF HAVING F7.2, 7H PPM = F7.4)
   70 CONTINUE
      WRITE (6,103) SMEAN, SVAR
  103 FORMAT (8HOMEAN = F7.2, 13H VARIANCE = F9.4/)
       WRITE (6,104) UP(N-1), SUM1
  104 FORMAT (33H PROBABILITY OF HAVING LESS THAN F7.2,
     1 7H PPM = F7.4 )
      WRITE (6,104) UP(N), SUM2
   12 CONTINUE
      STOP
      END
```

PROGRAM 2

```
C     PROGRAM FOR FINDING PROB. DISTRIBUTIONS WITH SMALL DO DEFICIT
      DIMENSION D(1000), A(1000), C(1000), T(1000), UP(1000),
     1 E(1000), B(1000), S(1000)
      READ (5.7) ZK1, ZK2, ZK3, ZLA, ZDB, DELTA, ALPHA, IP, NP
    7 FORMAT (7F6.0, 2I4)
      DNP = NP
      AIP = IP
      SK13 = ZK1 + ZK3
      F = ZLA / (DELTA * SK13)
      U = EXP(-F)
      JF = F
      NA = IP + JF
      CALL EXPON (NA, F, NA + 1, D)
      G = ZLA*ZK1/(DELTA*ZK2*SK13) + ZDB/(DELTA*ZK2)
      Y = EXP(-G)
      CALL EXPON (NP, G, NP+1, E)
      QT = ZK1/(ZK2-SK13)
      DO 14 NT = 1, 5
      TIME = NT
      PR = SK13 * TIME
      X = EXP (-PR)
      W = 1.0 - X
      CALL BINOM (W, IP, X, IP + 1, A, C)
      CALL SERMU (NA, IP, A, IP + 1, D, NA + 1, T, NA + 1)
      NAP1 = NA + 1
      DO 17 J = 1, NAP1
      BNP = J - 1
      UP(J) = BNP * DELTA
   17 T(J) = U * T(J)
      SMEAN = (F + AIP * X) * DELTA
      SVAR = DELTA * (ZLA / SK13 + AIP * DELTA * X * W )
      K = 1
      SUM1 = 1.0 - T(1)
    8 IF(SUM1 - ALPHA) 10, 10, 9
    9 K = K + 1
      SUM1 = SUM1 - T(K)
      GO TO 8
   10 SUM2 = SUM1 + T(K)
      WRITE (6,500) NT
  500 FORMAT (26HOTIME OF TRAVEL IN DAYS = I3/)
      WRITE (6,600)
  600 FORMAT (41H PROBABILITY MASS DENSITIES FOR POLLUTION)
      DO 60 KN = 1, NAP1
      IF (T(KN) - 0.00001) 60, 60, 80
   80 WRITE (6,601) UP(KN), T(KN)
  601 FORMAT (23H PROBABILITY OF HAVING F7.2, 7H PPM = F7.4)
   60 CONTINUE
      WRITE (6,700)SMEAN, SVAR
  700 FORMAT (8HOMEAN = F7.2, 13 H VARIANCE = F9.4/)
      WRITE (6,800) UP(K), SUM1
  800 FORMAT (33H PROBABILITY OF HAVING MORE THAN F7.2,7H PPM = F7.4)
      WRITE (6,800) UP(K-1), SUM2
      SR = ZK2 * TIME
      Z = EXP(-SR)
      COE = QT * (X-Z)
      WZ = 1.0 - COE
      CALL BINOM (WZ, IP, COE, IP+1, B, C)
      CALL SERMU (NP, IP, B, IP+1, E, NP+1, S, NP+1)
      NPP1 = NP + 1
```

```
      DO 16 NM = 1, NPP1
      ANP = NPP1 - NM
      UP(NM) = ANP * DELTA
   16 S(NM) = Y * S(NM)
      SMEAN = (DNP-G)*DELTA - AIP* DELTA* COE
      SVAR = DELTA*(G*DELTA + AIP*DEL1.*COE*(1.0- (X-Z)))
      KK = NP + 1
      SUM3 = S(KK)
   18 IF (ALPHA-SUM3) 20, 20,19
   19 KK = KK - 1
      SUM3 = SUM3 + S(KK)
      GO TO 18
   20 SUM4 = SUM3 - S(KK)
      WRITE (6,500) NT
      WRITE (6,101)
  101 FORMAT (38H PROBABILITY MASS DENSITIES FOR OXYGEN)
      DO 70 JK = 1, NPP1
      IF (S(JK) - 0.00001) 70, 70, 90
   90 WRITE (6,601) UP(JK), S(JK)
   70 CONTINUE
      WRITE (6,700) SMEAN, SVAR
      WRITE (6,104) UP(KK-1), SUM3
  104 FORMAT (33H PROBABILITY OF HAVING LESS THAN F7.2, 7H PPM=F7.4)
      WRITE (6,104) UP(KK), SUM4
   14 CONTINUE
      STOP
      END
```

PROGRAM 3

```
C     PROGRAM FOR FINDING PROB. DISTRIBUTIONS WITH VARIABLE
C     INITIAL CONDITIONS
      DIMENSION A(500), B(500), C(500), D(500), E(500), G(500),
     1 S(500), T(500), UP(500)
      READ (5.7) ZK1, ZK2, ZK3, ZLA, ZDB, DELTA, ALPHA, I, NP,
     1 J, MR2, MR1, MD2, MD1
    7 FORMAT (7F6.0, 714)
      AI = I
      AMR1 = MR1
      AMR2 = MR2
      MR21 = MR 2 - MR1
      P = (AI - AMR1) / (AMR2 - AMR1)
      SK13 = ZK1 + ZK3
      BP = ZLA / (DELTA * SK13)
      CONS = EXP(-BP)
      AJ = J
      ANP = NP
      D1 = MD1
      D2 = MD2
      R1 = MR1
      R2 = MR2
      MD21 = MD2 - MD1
      DSK = ZK2 - SK13
      TBP = (BP * ZK1) / DSK
      SBP = (ZLA * ZK1) / (DELTA * ZK2 * DSK)
      RBP = ZDB / (DELTA * ZK2)
      PP = (AJ - ZD1) / (ZD2 - ZD1)
      QP = 1.0 - PP
      ESK = ZK1 / DSK
      DO 12 NT = 1,5
      TIME = NT
      Z = SK13 * TIME
      U = EXP(-Z)
      TU = P * U
      W = BP* U
      SS = EXP(W)
      X = CONS * SS
      V = 1.0 - U
      TV = 1.0 - TU
      TW = BP * V
      CALL BINOM (V, MR1, U, MR1 + 1, B, E)
      CALL BINOM (TV, MR21, TU, MR21 + 1, C, D)
      CALL EXPON (MR2, TW, MR2 + 1, A)
       CALL SERMU (MR2, MR1, B, MR1+1, A, MR2+1, S, MR2+1)
      CALL SERMU (MR2, MR21, C, MR21 + 1, S, MR2 + 1, T, MR2 + 1)
      MRP1 = MR2 + 1
      DO 14 N = 1, MRP1
      BNP = N - 1
      UP(N) = BNP * DELTA
   14 T(N) = X * T(N)
      SMEAN = TW + DELTA * AI * U
      SVAR = DELTA * (TW + DELTA * AMR1 * U * V + DELTA *(AI-AMR1)*
     1 U * TV)
      N = 1
      SUM1 = T(1)
    8 IF (1.0 - ALPHA - SUM1) 10, 10, 9
    9 N = N + 1
      SUM1 = SUM1 + T(N)
```

```
    GO TO 8
 10 SUM2 = SUM1 - T(N)
    SUM3 = 1.0 - SUM1
    SUM4 = 1.0 - SUM2
    WRITE (6,500) NT
500 FORMAT (26H0TIME OF TRAVEL IN DAYS = I3/)
    WRITE (6,600)
600 FORMAT (41H PROBABILITY MASS DENSITIES FOR POLLUTION)
    DO 60 KN = 1, MRP1
    IF (T(KN) - 0.00001) 60, 60, 80
 80 WRITE (6,601) UP(KN), T(KN)
601 FORMAT (23H PROBABILITY OF HAVING F7.2, 7H PPM = F7.4)
 60 CONTINUE
    WRITE (6,700) SMEAN, SVAR
700 FORMAT (8H0MEAN = F7.2, 13H VARIANCE = F9.4/)
    WRITE (6,800) UP(N), SUM3
800 FORMAT (33H PROBABILITY OF HAVING MORE THAN F7.2, 7H PPM=F7.4)
    WRITE (6,800) UP(N-1), SUM4
    STU = -ZK2 * TIME
    UTN = EXP (STU)
    VTN = 1.0 - UTN
    CC = TBP * V + (RBP - SBP) * VTN
    EC = EXP (-CC)
    NP1 = NP + 1
    NPD = NP - MD2
    UAT = ESK * (U - UTN)
    VAT = 1.0 - UAT
    USA = P * UAT
    VSA = 1.0 - USA
    UUA = QP * UTN
    VVA = 1.0 - UUA
    CALL EXPON (NP, CC, NP + 1, A)
    CALL BINOM (VTN, NPD, UTN, NPD+1, B, C)
    CALL BINOM (VAT, MR1, UAT, MR1+1, D, C)
    CALL BINOM (VSA, MR21, USA, MR21+1, E, C)
    CALL BINOM (VVA, MD21, UUA, MD21 +1, G, C)
    CALL SERMU (NP, NPD, B, NPD+1, A, NP1, S, NP1)
    CALL SERMU (NP, MR1, D, MR1+1, S, NP1, T, NP1)
    CALL SERMU (NP, MR21, E, MR21 +1, T, NP1, S, NP1)
    CALL SERMU (NP, MD21, G, MD21+1, S, NP1, T, NP1)
    DO 110 N = 1, NP1
    BNP = NP1 - N
    UP(N) = BNP * DELTA
110 T(N) = EC * T(N)
    SMEAN = DELTA * (ANP + SBP * VTN - TBP * V - RBP * VTN
   1 - AI * UAT - (ANP - AJ) * UTN)
    SVAR = DELTA ** 2 * (TBP * V - SBP * VTN + RBP * VTN
   1 + (ANP-D2) * UTN * VTN + R1 * UAT * VAT +(R2 - R1) * USA
   2 * VSA + (D2 - D1) * UUA * VVA)
    SUM1 = T(N)
 18 IF (ALPHA - SUM1) 20, 20, 19
 19 N = N - 1
    SUM1 = SUM1 + T(N)
    GO TO 18
 20 SUM2 = SUM1 - T(N)
    WRITE (6,500) NT
    WRITE (6,101)
101 FORMAT (38H PROBABILITY MASS DENSITIES FOR OXYGEN)
    DO 90 JK = 1, NP1
    IF (T(JK) - 0.00001) 90, 90, 70
```

PROGRAM 3 CONTD

```
 70 WRITE (6,601) UP(JK), T(JK)
 90 CONTINUE
    WRITE (6,700) SMEAN, SVAR
    WRITE (6,102) UP(N-1), SUM1
102 FORMAT (33H PROBABILITY OF HAVING LESS THAN F7.2, 7HPPM = F7.4)
    WRITE (6,102) UP(N), SUM2
 12 CONTINUE
    STOP
    END
```

SUBROUTINE 1

```
    SUBROUTINE EXPON (I, W, M, A)
    DIMENSION A(M)
    A(1) = 1.0
    DO 10 K = 1, I
    KP1 = K + 1
    CK = K
 10 A(KP1) = A(K) * W / CK
    RETURN
    END
```

SUBROUTINE 2

```
    SUBROUTINE BINOM (V, I, U, M, B, C)
    DIMENSION B(M), C(M)
    C(1) = 1.0
    B(1) = V ** I
    DO 8 J = 1, I
    JP1 = J + 1
    AI = I
    AJ = J
    AJM1 = J - 1
    C(JP1) = C(J) * (AI - AJM1) * U / AJ
  8 B(JP1) = C(JP1) * (V **(I -J))
    RETURN
    END
```

SUBROUTINE 3

```
   SUBROUTINE SERMU (NP, I, D, MN, S, NN, T, KK)
   DIMENSION S(NN), D(MN), T(KK)
   NP1 = NP + 1
   IF (NP - 1) 13, 13, 15
13 DO 5 M = 1, NP1
   T(M) = 0.0
   DO 5 L = 1, M
   MIL = M + 1 - L
 5 T(M) = S(L) * D(MIL) + T(M)
   RETURN
15 IP1 = I + 1
   DO 22 N = 1, IP1
   T(N) = 0.00
   DO 22 L = 1, N
   NIL = N + 1 - L
22 T(N) = S(L) * D(NIL) + T(N)
   IP2 = IP1 + 1
   DO 24 K = IP2, NP1
   T(K) = 0.00
   I2 = K - I
   DO 24 M = I2, K
   KM1 = K - M + 1
24 T(K) = S(M) * D(KM1) + T(K)
   RETURN
   END
```

APPENDIX II - MATHEMATICS

PROBABILITY GENERATING FUNCTIONS

The first part of this section will be devoted to the study of some of the properties of generating functions. The probability generating function for two random variables is defined by

$$P(S,R;t) = \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N}(t) S^M R^N$$

where $Pr_{M,N}(t)$ is the joint probability that the pollution is in state M and the dissolved oxygen is in state N. Given the joint probability mass densities, one can multiply by S^M and R^N and sum to obtain the generating function. Also if one is given P(S,R;t), this can be expanded in a power series in R and S whose coefficients will be the joint probability mass densities. Similar statements are true if one is considering one of the two marginals.

If R = 1, the joint density is summed over N and the marginal probability generating function is formed. A similar result is obtained if S = 1.

The term $\sum_{M=0}^{\infty} P_{M,O}(t)\, S^M$ can be written as P(S,O;t). This can easily be shown by letting R = 0 when P(S,R;t) is written in summation notation.

Consider next some of the algebraic manipulations involved in writing an equation in terms of generating functions.

1. $$\sum_{M=0}^{\infty} \sum_{N=0}^{\infty} M Pr_{M,N}(t) S^M R^N = S \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N}(t)\, M S^{M-1} R^N$$

$$= S \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N}(t)\, \frac{\partial S^M}{\partial S}\, R^N = S\, \frac{P}{S}.$$

2. $\sum_{M=1}^{\infty} \sum_{N=0}^{\infty} Pr_{M-1,N}(t)\ S^M R^N = \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N}(t)\ S^{M+1} R^N = S\ P.$

In going from the first term to the second term, M was replaced by M + 1.

3. $\sum_{M=0}^{\infty} \sum_{N=0}^{\infty} (M+1)\ Pr_{M+1,N}(t)\ S^M R^N = \sum_{M=1}^{\infty} \sum_{N=0}^{\infty} M\ Pr_{M,N}(t)$

$S^{M-1} R^N = \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N}(t)\ \frac{\partial S^M}{\partial S}\ R^N = \frac{\partial P}{\partial S}$. In going from the first term to the second M+1 was replaced by M. The index on M could be lowered to zero because the value of the term added is zero.

4. $\sum_{M=0}^{\infty} \sum_{N=0}^{\infty} (M+1)\ Pr_{M+1,\ N+1}(t)\ S^M R^N = \frac{\partial}{\partial S} \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N+1}(t)$

$S^M R^N = \frac{\partial}{\partial S} \sum_{M=0}^{\infty} \sum_{N=1}^{\infty} Pr_{M,N}(t) S^M R^{N-1} = \frac{1}{R} \frac{\partial}{\partial S} \sum_{M=0}^{\infty} (\sum_{N=0}^{\infty}$

$Pr_{M,N}(t) S^M R^N - Pr_{M,O} S^M) = \frac{1}{R} \frac{\partial P}{\partial S} - \frac{1}{R} \frac{\partial}{\partial S} P(S,O;t)$. The change from the first term to the second is the same as Number 3. In going from the second to the third, N was replaced by N-1.

5. $\sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N+1}(t) S^M R^N = \frac{1}{R}\ P - \frac{1}{R}\ P(S,O;t).$

The changes involved here are the same as those involved in the latter part of Number 4.

6. $\sum_{M=0}^{\infty} \sum_{N=1}^{\infty} Pr_{M,N-1}(t) S^M R^N = R\ P.$

This is similar to Number 2 above.

7. $\sum_{M=0}^{\infty} \sum_{N=1}^{\infty} (N-1) Pr_{M,N-1}(t) S^M R^N = R^2 \frac{\partial}{\partial R} \sum_{M=0}^{\infty} \sum_{N=1}^{\infty} Pr_{M,N-1}(t)$

$S^M R^{N-1} = R^2 \frac{\partial}{\partial R} \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} Pr_{M,N}(t) S^M R^N = R^2 \frac{\partial P}{\partial R}$.

In going from the second to the third term N-1 was replaced by N.

PARTIAL DIFFERENTIAL EQUATION

The Stochastic development requires the solution of a partial differential equation. Some of the mathematics involved will be given here. One possible differential equation reference is the book by Nelson, Folley, and Coral (9). It is desired to find a solution to a partial differential equation of the form

$$A_1 \frac{\partial Y}{\partial x_1} + A_2 \frac{\partial Y}{\partial x_2} + A_3 \frac{\partial Y}{\partial x_3} = C \qquad \text{(II.1)}$$

where A_1, A_2, A_3 and C are functions of the dependent variable y and independent variables x_1, x_2, and x_3. This equation is a quasi-linear partial differential equation of the first order.

Consider also the linear partial differential equation of the first order

$$A_1 \frac{\partial \phi}{\partial x_1} + A_2 \frac{\partial \phi}{\partial x_2} + A_3 \frac{\partial \phi}{\partial x_3} + C \frac{\partial \phi}{\partial y} = 0 \qquad \text{(II.2)}$$

which will have for its solution $(x_1, x_2, x_3, y) = a$, where a is the constant of integration.

When $(x_1, x_2, x_3, y) = a$ is solved for y in terms of x_1, x_2, x_3, and a, it can be shown that the result will be a solution for (II.1). Hence if a solution is found for (II.2) a solution for (II.1) is easily obtained.

The solution of (II.2) can be found using Lagrange's Method, which is outlined in the following paragraphs.

Consider the system of ordinary differential equations (called subsidiary equations)

$$\frac{dx_1}{A_1} = \frac{dx_2}{A_3} = \frac{dx_3}{A_3} = \frac{dy}{C} \quad . \qquad \text{(II.3)}$$

It can be shown that any solution of the subsidiary equation is also a solution of the partial differential equation (II.2). Now the problem has been reduced to solving a seried os ordinary differential equations (II.3). For this case choose three independent equations, which are integrable, from the subsidiary equations. Perform the integrations and obtain the solutions.

$$\phi_1(x_1,x_2, x_3,y) = c_1 \qquad \phi_2(x_1,x_2.x_3,y) = c_2$$

$$\phi_3(x_1,x_2,x_3.y) = c_3 \tag{II.4}$$

where c_1,c_2, and c_3 are the constants of integration. The above equations can each be solved for y and each of the y's will be a general solution of the original partial differential equation (II.1).

Any assumed initial condition will determine a unique functional relationship among ϕ_1, ϕ_2, and ϕ_3. Any arbitrary function $\theta(\phi_1, \phi_2,\phi_3) = 0$ possessing partial derivatives will be a general solution of (II.2). This arbitrary function can also be written

$$\psi(\phi_1, \phi_2) = \phi_3, \tag{II.5}$$

this latter form being more convenient. The ψ function can be explicitly determined if initial conditions are given. The initial condition is a given relationship among the variables x_1,x_2, x_3 and y.

Suppose the initial condition is

$$y = \rho(0,x_2,x_3) \tag{II.6}$$

In (II.5) make the substitution of (II.6), that is, let $x_1=0$ and replace y by the given function of x_2 and x_3. This substitution means that (II.5) must hold for the initial conditions. Hence

$$\psi(\phi_1(0,x_2,x_3,\ \rho(0,x_2,x_3),\ {}_2(0,x_2,x_3,\ \rho(0,x_2,x_3,)))$$

$$= \phi_3(0,x_2,x_3,\ \rho(0,x_2,x_3,)) \qquad \text{(II.7)}$$

Let $v = \phi_1(0,x_2,x_3,\ \rho(0,x_2,x_3,))$ and $w = \phi_2(0,x_2,x_3,\ \rho(0,x_2,x_3,),$ (II.8)

then solve these two equations for x_2 and x_3 in terms of v and w.

Then this result is substituted into (II.7) to obtain $\psi(v,w)$ as a function of v and w. Now the form of the ψ function is known completely and this has been determined by the initial conditions.

In particular (II.5) is what is desired, so in (II.8) let $v= \phi_1$ and $w=\phi_2$. Then $\phi_3= \psi(\phi_1,\phi_2)$ is a known function of ϕ_1,ϕ_2 and ϕ_3. Replace ϕ_1, ϕ_2, and ϕ_3 by their equals in (II.4).

We can then solve for y as a function of x_1, x_2, and x_3. Then the partial differential equation is satisfied because the result is a function of the ϕ's and the initial condition is satisfied because it determined the form of ψ .

MOMENT GENERATING FUNCTIONS

A joint moment generating function is defined as

$$M(\theta,\phi;t) = \sum_{M=0}^{\infty} \sum_{N=0}^{\infty} P_{M,N}(t)e^{\theta M}e^{\phi N}. \qquad \text{(II.)}$$

If S is replaced by e^{θ} and R is replaced by e^{ϕ} in the probability generating function, one obtains the moment generating function

$$P(e^{\theta}\ ,\ e^{\phi};t) = M(\theta,\phi"t).$$

The marginal moment generating function $M(\theta,t)$ is found by setting $\phi=0$ in $M(\theta,\phi;t)$. It can also be found by replacing S by e^{θ} in the marginal probability generating function $P(e^{\theta},1;t)$

$$M(\theta,t) = \sum_{M=0}^{\infty} P_M(t)e^{\theta M} \qquad \text{(II.10)}$$

The first moment, i.e., the mean, is found by differentiating the moment generating function (II.10) with respect to θ and evaluating the result at $\theta=0$.

$$\mu_1 = \frac{M(\theta,t)}{\partial\theta}\bigg|_{\theta=0} = \sum_{M=0}^{\infty} MP_M(t) \qquad \text{(II.11)}$$

The second non-central moment is the second partial derivative of $M(\theta,t)$ with respect to θ, evaluated at $\theta=0$.

$$\mu_2 = \frac{\partial^2 M(\theta,t)}{\partial\theta^2}\bigg|_{\theta=0} = \sum_{M=0}^{\infty} M^2 P_M(t)$$

The variance is then $\mu_2 - \mu_1^2$

$$\mu_{2'} = \frac{\partial^2 M(\theta,t)}{\partial\theta^2}\bigg|_{\theta=0} - \left(\frac{\partial M(\theta,t)}{\partial\theta}\bigg|_{\theta=0}\right)^2. \qquad \text{(II.12)}$$

Other non-central and central moments can be found similarly.

APPENDIX III - METHODS OF MEASUREMENT

CHEMICAL OXYGEN DEMAND

The COD procedure is included in _Standard Methods for The Examination of Water and Waste Water_ (1). The first step is to obtain 50 ml. of settled sample. The settled sample was used because the sugar, the amount of which is what is actually being measured, is in solution, and it is possible that solids such as small pieces of grass in the pond water, if not settled out, could be oxidized during the COD test. These solids would interfere with the pollution measurement because the bacteria would not attack them at the same rate at which the sugar is attacked and samples would not contain a uniform amount of them.

To this sample is added potassium dichromate and sulfuric acid. The mixture is refluxed for two hours while the sulfuric and chromic acids oxidize the sugar and any other oxidizable material. After cooling, the excess potassium dichromate is titrated with ferrous ammonium sulfate in the presence of ferroin indicator. The more pollution in the sample the smaller the amount of ferrous ammonium sulfate needed to arrive at the endpoint, this endpoint representing change in color. A blank of 50 ml. of distilled water is run as described above. The blank, having no pollution, will always use more ferrous ammonium sulfate than the sample. Then the normality of the ferrous ammonium sulfate is ascertained by standardizing it against the potassium dichromate. Finally, the COD can be calculated using

$$\text{mg/liter COD} = \frac{(a-b)c \times 8{,}000}{50 \text{ ml sample}} \quad \text{where}$$

a = ml of ferrous ammonium sulfate used for blank

b = ml of ferrous ammonium sulfate used for sample

c = calculated normality of ferrous ammonium sulfate.

Several points of this test require further discussion. The test has two features which increase the accuracy. First, the end point of the titration is not exact, so the blank is titrated first and an endpoint color is established. Then the samples are titrated; the final color of the sample is titrated to correspond to the color of the blank. Thus, when the subtraction is made in the formula, the effect of one drop more or less in both the blank and sample is eliminated. The second aid to increased accuracy is the standardization of the ferrous ammonium sulfate each time the test is run. This eliminates the effect of slight errors in the initial concentration of the ferrous ammonium sulfate and also the effect of any time change in the concentration.

Better results can also be obtained if the concentrations are chosen so that (a-b) is about 1/2 a. For this the textbook recommended that dilute potassium dichromate be used, .025N, and the ferrous ammonium sulfate concentration was then chosen to be 1/12 N and toward the end of the run 1/24 N was used. If the ferrous ammonium sulfate is too dilute the effect of one drop is not noticed in the titration.

This test is best used for high concentrations of pollution. When the dilute potassium dichromate is used, any small amount of dirt on glassware can be oxidized and may greatly influence the test. This is especially true with the blank. Even small amounts of stopcock grease on the reflux apparatus can influence the test. The manual states that whenever concentrations of pollution fall below 20 ppm, the test is accurate only to an order of magnitude of 1 ppm.

BIOCHEMICAL OXYGEN DEMAND

In order to compare the COD pollution measurement with the standard BOD several BOD measurements were made during the run. For this test, settled samples of 5 and 10 ml were taken from each can initially, and after the pollution decreased, samples of 10 and 20 ml were taken. These samples were placed in standard BOD bottles and the bottles were filled with dilution water. The dilution water was distilled water containing small quantities of phosphate buffer, magnesium sulfate, calcium chloride and ferric chloride. The bottles were sealed and stored for five days. Enough dilution water was made so that there was some remaining and it was used to determine the initial dissolved oxygen content of the sealed bottles. Dissolved oxygen concentration is measured using the Winkler Method to be discussed later in this section. After five days the 10 bottles were opened and their dissolved oxygen content determined. The 5-day BOD is then determined as follows:

$$\text{mg/liter BOD} = \frac{D_0 - D_2}{P} \qquad \text{where}$$

D_0 = DO of original dilution water

D_2 = DO of diluted sample after incubation

P = ml sample/ml in BOD bottle.

For this test to give accurate results the difference D_0-D_2 must be greater than one ppm and there must be a residual concentration of one ppm. A difference which is too small is greatly affected by slight errors in measuring D_0 and D_2. If D_2 is too small then the bacteria have slowed down due to a lack of oxygen. In order to insure that at least one result would be in the desired range, two dilutions were made.

It is now evident why the COD was preferable to the BOD. The COD takes five hours while the BOD takes five days. The BOD requires many bottles for storage, requires time to make the dilution water, and requires care that the dilution water is not aerated in transferring it between vessels.

DISSOLVED OXYGEN

Two methods were used to make the dissolved oxygen determinations. The first is a chemical method which is known as the Winkler Method. A 132 ml sample was collected in a glass stoppered bottle and manganous sulfate and alkali-iodide-azide reagent were added. The resultant manganese hydroxide floc is shaken and allowed to settle twice. Then with the addition of sulfuric acid, iodine is liberated. The amount now taken for titration is equivalent to 100 ml of the original sample, i.e.,

$$100 \times \frac{132}{132-2} = 101.5 \text{ ml}$$

where the 2 is the 2 ml of sample displaced by the manganese sulfate and the alkali-iodide-azide reagent. Finally, the 101.5 ml is titrated with sodium thiosulfate using starch as the indicator. The end point of the titration is the disappearance of the blue color. One ml of 0.025 N sodium thiosulfate is equivalent to 0.2 mg in DO. For a volume equal to 100 ml of original sample

1 ml thiosulfate = concentration of 2 mg/liter of dissolved oxygen.

Hence, the amount of thiosulfate used is multiplied by 2 to obtain the concentration of dissolved oxygen in mg/liter or ppm. It is necessary to standardize the thiosulfate regularly with either potassium biniodate or potassium dichromate and make any necessary adjustments in the thiosulfate concentration to bring it to 0.025N.

The other dissolved oxygen determination is made using a galvanic cell oxygen analyser. The probe is held in the can approximately two inches below the surface, two inches from the mechanical stirrer shaft and pointed in the direction opposing the flow. The meter readings are affected by the velocity of the water passing the probe. It is essential therefore to place the probe in the same position each time. The meter stabilizes after approximately two minutes and the meter is read to the nearest 0.05 units. The meter must be calibrated and this is done as follows: A sample is taken from one can and a Winkler determination is made of the dissolved oxygen. The calibration factor ϕ is defined as

$$\phi = \frac{\text{Meter DO}}{\text{Winkler Method DO}} .$$

All other meter readings are divided by ϕ to obtain the actual dissolved oxygen. No temperature correction is needed as all work was done in a constant temperature room at 21°C.

CHLORIDE

The Mohr Method for chloride is included in Standard Methods for the Examination of Water and Waste Water (1). A 100 ml sample is taken and 1 ml of potassium chromate is added as an indicator. This is titrated with silver nitrate to a pinkish yellow endpoint. A blank of 100 ml of distilled water is titrated to the same color. Also, the silver nitrate must be standardized against a known solution of sodium chloride. Then

$$\text{mg/l Cl} = \frac{(a-b)c \times 35{,}460}{100 \text{ ml sample}}$$ where

a = ml silver nitrate for the sample

b = ml silver nitrate for the blank

c = normality of silver nitrate.

REFERENCES

1. American Public Health Association and American Water Works Association. *Standard Methods for the Examination of Water and Sewage*, 11th ed. New York: John Wiley & Sons, 1961.

2. Bailey, N. T. J. *The Elements of Stochastic Processes*. New York: John Wiley & Sons, 1964.

3. Bharucha-Reid, A. T. *Elements of the Theory of Markov Processes and Their Applications*. New York: McGraw-Hill, 1960.

4. Churchill, M. A. and Buckingham, R. A. "Statistical Method for Analysis of Stream Purification Capacity," *Sewage and Industrial Wastes*, Vol. 28 (1956). pp. 517-526.

5. Dobbins, W. E. "BOD and Oxygen Relationships in Streams," *Jour. Sanitary Engineering Div. Proc. ASCE*, Vol. 90 (1964), pp. 53-77.

6. Fair, G. M., and Geyer, J. C. *Elements of Water Supply and Waste-Water Disposal*. New York: John Wiley & Sons, 1961.

7. Feller, W. *An Introduction to Probability Theory and Its Applications*, 2nd ed. New York: John Wiley & Sons, 1961.

8. Kells, L. M. *Elementary Differential Equations*. New York: McGraw-Hill, 1954.

9. Nelson, A. L., Folley, K. W., and Coral, M. *Differential Equations*. Boston: D. C. Heath and Company, 1960.

10. Resources Agency of California Department of Water Resources: *Sacramento River Water Pollution Survey*, Bulletin No. 111 and Appendix B (August, 1962).

11. Streeter, H. W. and Phelps, E. B. "A Study of the Pollution and Natural Purification of the Ohio River," *Public Health Bulletin 146*, U. S. Public Health Service, Washington, D. C., 1925.

12. Thomann, R. "Mathematical Model for Dissolved Oxygen," *Jour. Sanitary Engineering Div. Proc. ASCE*, Vol. 89 (1963). pp. 1-30.

13. White, J. M. "Pollution and America's Rivers," *The Washington Post*. May, 1965.

bullet

water resources research center
virginia polytechnic institute blacksburg