FINAL REPORT

A PROBABILISTIC METHODOLOGY FOR ANALYZING WATER QUALITY EFFECTS OF URBAN RUNOFF ON RIVERS AND STREAMS

JULY, 1989

OFFICE OF WATER
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

DISCLAIMER

This report has been reviewed by the U.S. Environmental Protection Agency and approved for release. Approval does not signify that the contents necessarily reflect any policies or decisions of the U.S. Environmental Protection Agency or any of its offices, grantees, contractors, or subcontractors.

ABSTRACT

The Probability Dilution Model provides a screening level methodology for determining the effect of intermittent pollutant discharges on the water quality of streams and rivers. The processes analyzed in the model (urban runoff, streamflow, and water quality) are probabilistic by nature and are treated as such by the model. The method uses a determinisitic mass balance approach and accounts for the natural variabilities in the model inputs to estimate a cumulative probability distribution of in-stream concentration due to runoff into the stream. This model is not meant to replace more sophisticated deterministic or probabilistic models, but is an extremely useful screening level tool for initial planning purposes to help determine problem areas and to screen possible control alternatives. Three methods, the Analytically Derived Distribution Approach, a numerical Gaussian Quadrature method and the Monte Carlo simulation approach, are presented.

The derived distribution method is based on analytically estimating the mean and variance of the downstream concentration assuming that the upstream flow, runoff, and downstream flow as well as pollutant concentrations are lognormally distributed. The method is conducive to hand calculations.

The Gaussian Quadrature Method assumes that the upstream flow and concentration and the runoff flow and concentration are lognormally distributed but makes no assumptions about the downstream flow and concentration. This method is more accurate than the previous method, but requires more computational effort.

The Monte Carlo Simulation Approach allows flexibility in specifying the input distributions. This method does not make any assumptions about the downgradient distributions, is easy to understand and general in its approach. However, it requires the repeated application (500 to 1000 times) of the deterministic mass balance approach and hence requires more computational effort.

The Probability Dilution Model is easy to implement and use. A computer code has been written in FORTRAN 77 and tested on IBM-compatible computers. Comparison of results obtained using the three approaches discussed above using identical input parameters is also presented. A copy of the source code and an executable code compiled using the Ryan-McFarland compiler is included along with sample input and output files.

TABLE OF CONTENTS

CHAPTER		PAGE
	Disclaimer	i
	Abstract	ii
	Table of Contents	iii
	List of Tables	v
	List of Figures	vi
	Acknowledgements	viii
1.0	INTRODUCTION	1-1
1.1	Background	1-1
1.2	Nonpoint Source Technical Assistance	1-2
1.3	Organization of this Report	1-2
2.0	TECHNICAL OVERVIEW	2-1
2.1	Introduction	2-1
2.2	The System Being Simulated	2-1
2.3	Relevant Management Decision	2-1
2.4	Classification of Models	2-2
2.4.1	Approaches to Runoff Modeling	2-3
2.4.2	Regression Models	2-3
2.4.3	Simulation Models	2-4
2.4.4	Probabilistic Methods	2-4
2.5	Characterization of Rainfall and Runoff Events	2-5
2.5.1	The Rainfall/Runoff Event	2-5
2.5.2	Selection of Interevent Duration	2-5
2.5.3	Estimation of Runoff Statistics from Rainfall Statistics	2-6
2.5.3.1	Quantity Statistics	2-6
2.5.3.2	Quality Statistics	2-6
2.6	Summary	2-7
3.0	THEORETICAL BASIS OF MODEL	3-1
3.1	Introduction	3-1
3.2	Model Framework and Statistical Concepts	3-1
3.3	The Analytically Derived Distribution Approach	3-3
3.4	Key Assumptions for the Analytically Derived Distribution Approach	3-7
3.5	The Direct Numerical Solution	3-7
3.6	Key Assumptions in the Direct Numerical Solution	3-11
3.7	The Monte Carlo Simulation Approach	3-11
3.8	Key Assumptions of the Monte Carlo Simulation Approach	3-12
3.9	Recurrence Interval Calculation	3-12

TABLE OF CONTENTS (concluded)

CHAPTER		PAGE
4.0	APPLICATION OF METHODOLOGY	4-1
4.1	Introduction	4-1
4.2	Data Needs and Sources	4-1
4.3	Computation Procedure Using the Analytically Derived Distribution Approach	4-3
4.4	Illustrative Example of Analysis	4-7
4.5	Comparison of the Three Probabilistic Methodologies	4-8
4.6	Interpretation and Evaluation of Results	4-13
5.0	EVALUATION OF METHODOLOGY	5-1
5.1	Introduction	5-1
5.2	Verification of Computed Concentrations	5-1
5.3	Verification of Implied Effect on Beneficial Use	5-5
5.4	Sensitivity Analyses	5-7
5.4.1	Analytically Derived Distribution Approach Vs. Direct Numberical Approach	5-7
5.4.2	Correlation between Streamflow and Runoff Flow	5-12
5.4.3	Effect of Upstream Concentrations	5-12
5.5	Summary and Discussion	5-15
6.0	USER'S GUIDE	6-1
6.1	Model Acquisition and Installation Procedures	6-1
6.1.1	Model Acquisition	6-1
6.1.2	Model Installation	6-1
6.1.3	Testing the Model	6-2
6.2	Model Structure	6-2
6.3	Common Blocks and Parameter Statements	6-2
6.4	Model Input	6-7
6.4.1	Analytic Model Data Requirements	6-7
6.4.2	Monte Carlo Data Requirements	6-11
6.5	Example Application	6-18
	REFERENCES	R-1
	NOTATIONS	N-1
	APPENDIX A : Mathematical Details of the Numerical Procedure	A-1
	APPENDIX B: Example Application of the Method of Moments Method	B-1
	APPENDIX C: List of Subroutines Included in the PDM Model	C-1
	APPENDIX D: List of Variables Included in the PDM Model	D-1

LIST OF TABLES

TABLE		PAGE
4-1	INPUT DATA REQUIREMENTS AND RELATIONSHIPS FOR LOGNORMAL DISTRIBUTIONS	4-2
4-2	INPUTS USED IN SENSITIVITY ANALYSIS OF PDM MODEL	4-14
4-3	COMPARISON OF RESULTS BETWEEN MONTE CARLO METHOD, ANALYTICALLY DERIVED DISTRIBUTION APPROACH, AND DIRECT NUMERICAL APPROACH	4-15
5-1	RAPID CITY VERIFICATION TEST	5-3
6-1	INPUT FORMATS FOR ANALYTICAL PDM MODEL	6-8
6-2	SAMPLE INPUT FILE PDM1.IN FOR ANALYTIC DAM MODELS	6-10
6-3	SAMPLE OUTPUT FILE PDM1.OUT	6-12
6-4	FORMAT FOR READING SIMULATION CONTROL PARAMETERS (DATA GROUP 1)	6-13
6-5	FORMAT FOR READING DISTRIBUTION PARAMETERS FOR INPUT VARIABLES (DATA GROUP 2)	6-14
6-6	FORMAT FOR READING EMPIRICAL DISTRIBUTION DATA (DATA GROUP 3)	6-15
6-7	FORMAT FOR READING OUTPUT OPTIONS (DATA GROUP 4)	6-16
6-8	FORMAT FOR READING CORRELATED VARIABLES INPUT (DATA GROUP 5)	6-17
6-9	SAMPLE OUTPUT FILE PDM1.OUT	6-19
6-10	SAMPLE OUTPUT FILE PDM1MC.OUT	6-22

LIST OF FIGURES

FIGURE		PAGE
3-1	SCHEMATIC REPRESENTATION OF WATER QUALITY EFFECTS OF URBAN RUNOFF ON RIVERS AND STREAMS	3-2
3-2	REGIONS OF INTEGRATION FOR THE JOINT PROBABILITY DENSITY FUNCTION	3-9
4-1	SCHEMATIC OUTLINE OF PROBABILISTIC ANALYSIS METHOD	4-4
4-2	CONSTRUCTION OF PROBABILITY DISTRIBUTION PLOT	4-6
4-3	PROBABILITY DISTRIBUTION OF POLLUTANT CONCENTRATIONS DURING STORM RUNOFF EVENTS	4-9
4-4	RECURRENCE INTERVALS FOR STREAM CONCENTRATIONS	4-10
4-5a	COMPARISON OF CDFS FOR THREE PROBABILISTIC METHODOLOGIES	4-11
4-5b	COMPARISON OF CDFS FOR THREE PROBABILISTIC METHODOLOGIES (HIGHEST 10% OF VALUES)	4-12
4-6	USE OF THE METHODOLOGY TO EVALUATE THE WATER QUALITY IMPACT OF URBAN RUNOFF CONTROL PROGRAM (I)	4-16
4-7	USE OF THE METHODOLOGY TO EVALUATE THE WATER QUALITY IMPACT OF URBAN RUNOFF CONTROL PROGRAM (II)	4-17
5-1	MAP OF RAPID CREEK AND MAJOR DRAINAGE BASINS WITHIN RAPID CITY PROPER: SAMPLING SITE LOCATIONS	5-2
5-2	PROBABILITY DISTRIBUTION OF STREAM AND RUNOFF FLOWS	5-3
5-3	COMPARABLE INFORMATION ON CONCENTRATIONS IN RUNOFF AND UPSTREAM LOCATIONS AND THE PROBABILITY DISTRIBUTION	5-4
5-4	RESULTS FOR TOTAL P	5-4

LIST OF FIGURES (concluded)

FIGURE		PAGE
5-5	RESULTS FOR COD	5-6
5-6	RESULTS FOR LEAD	5-6
5-7	COMPARISION OF RESULTS FROM THE DERIVED DISTRIBUTION AND NUMERICAL SOLUTION APPROACH - PROBABILITY DISTRIBUTION FOR SELECTED INPUT CONDITIONS	5-8
5-8	COMPARISION OF RESULTS FROM THE DERIVED DISTRIBUTION AND DIRECT NUMERICAL SOLUTION APPROACHES - RECURRENCE INTERVALS FOR SELECTED INPUT CONDITIONS	5-9
5-9	DIFFERENCE IN RESULTS BETWEEN DERIVED DISTRIBUTION AND NUMERICAL SOLUTION APPROACH OVER A WIDE RANGE OF INPUTS	5-11
5-10	EFFECT OF STREAM/RUNOFF FLOW CORRELATIONS	5-13
5-11	EFFECT OF UPSTREAM CONCENTRATION ON MEAN RECURRENCE INTERVAL	5-14
5-12	EFFECT OF CONTROL WITH NO BACKGROUND CONTRIBUTION	5-16
5-13	EFFECT OF CONTROL WITH HIGH BACKGROUND CONTRIBUTION	5-17
6-1	FLOWCHART OF THE PDM MODEL	6-3
6-2	FLOW CHART OF THE ANALYTICAL PDM MODEL	6-4
B-1	EXAMPLE APPLICATION OF THE METHOD OF MOMENTS METHOD	B-1
B-2(a)	PROBABILITY DISTRIBUTION OF POLLUTANT CONCENTRATIONS DURING STORM RUNOFF EVENTS	B-7
B-2(b)	RECURRENCE INTERVALS FOR STREAM CONCENTRATIONS	B-7

ACKNOWLEDGEMENTS

This report was prepared by Eugene D. Driscoll (Woodward-Clyde Consultants), Philip E. Shelley (EG&G Washington Analytical Services Center, Inc.), David R. Gaboury, and Atul Salhotra (Woodward-Clyde Consultants). The PC-based computer program was developed by Phil Mineart and Atul Salhotra also of Woodward-Clyde Consultants. Dominic M. DiToro (HydroQual Inc.) conceived the approach, developed the methodology, and provided comment and advice on preparation of this manual. Much of Chapter 3 describing the theoretical basis of the model has been extracted from his original technical paper on the method. John L. Mancini (John Mancini Consultants) made significant contributions to the application of the methodology as presented herein.

Special thanks are due the members of the SAB's Probabilistic Methodology Subcommittee—B. C. Dysart, III, Chairman, R. A. Conway, R. C. Loehr, D. J. O'Connor, and C. R. O'Melia—and their consultants (B. Adams and M. Small) for their careful review of the draft report. Acknowledgements are also due to Mr. Tom Barnwell and Rudy Parish of the U.S. EPA Environmental Research Laboratory at Athens, Georgia and Tom Davenport of U.S. EPA Region VI for reviewing the draft report and providing many useful comments.

The efforts of Dennis N. Athayade of the Office of Water's Nonpoint Sources Branch are especially appreciated, both for providing the vision which led to the development of the methodology as well as serving as Project Officer.

Chapter 1.0

INTRODUCTION

1.1 BACKGROUND

This report presents a methodology for assessing the effects of intermitttent, highly variable discharges, such as urban runoff, on the water quality of rivers and streams. The methodology was developed under the sponsorship of the U.S. Environmental Protection Agency (U.S. EPA) over the past 15 years, through a number of contracts, and traces its roots to the statistical methods effort for urban stormwater initiated by the (then) Water Planning Division in 1974. The method was also discussed in the *Areawide Assessment Procedures Manual* published by the Office of Research and Development, Municipal Environmental Research Laboratory, in 1976. Much of the additional work that brought the methodology to its present form was supported by the Nonpoint Sources Branch of the Office of Water through its Nationwide Urban Runoff Program (NURP), which was concluded in early 1984, and subsequent work assignments under a series of nonpoint source technical assistance contracts.

The Nationwide Urban Runoff Program (NURP) was conducted by the EPA and many cooperating federal, state, regional, and local agencies, distributed widely across the United States. The program yielded a great deal of information, much of which will be useful for many years, for a broad spectrum of planning activities. Among the many tangible products of the program are the Final Report, the reports of various grantees (available from National Technical Information Service, NTIS), and several technical reports that focused on specialized aspects of the program, its techniques, and its findings.

One of the final NURP products was the first draft of a report entitled, "A Probabilistic Methodology for Analyzing Water Quality Effects on Rivers and Streams," dated February 15, 1984. In keeping with established NURP practice, the document was issued as a draft report to allow maximum opportunity for peer review and comment. Because of the perceived importance of the methodology and its potential for widespread application in conducting screening level analysis for nonpoint source impacts on rivers and streams, the Office of Water also requested the Environmental Engineering Committee of the Agency's Science Advisory Board (SAB) to provide a formal review of the draft report. This was done by an *ad hoc* Probabilistic Methodology Subcommittee, which held hearings on the subject and retained consultants for specialized advice. The collective comments resulting from this process were consolidated in a report that was sent to the Administrator on July 31, 1985. The letter forwarding the report to the Administrator stated that "The Committee believes strongly that statistically based approaches to water quality management are an important tool for the decision-maker, and commends the Agency for supporting the effort under review."

The present document represents the final revision of that earlier report and incorporates the SAB comments as well as those received from other reviewers. As a result of the extensive review process, the Agency considers the method to be technically sound and recommends its use as a screening level tool in dealing with urban runoff and similar nonpoint source discharges. Practitioners are urged to read this document

carefully, since the Agency does not believe that the method should be used by individuals who do not fully understand the approach and the assumptions inherent therein, nor should it be applied in situations for which it is not technically suitable.

1.2 NONPOINT SOURCE TECHNICAL ASSISTANCE

The possible deleterious water quality effects of nonpoint sources in general, and urban runoff in particular, were recognized by the Water Pollution Control Act Amendments of 1972. Because of uncertainties about the true significance of urban runoff as a contributor to receiving water quality problems, Congress made treatment of separate stormwater discharges ineligible for federal funding when it enacted the Clean Water Act in 1977. With the enactment of the Water Quality Act of 1987, Congress expressed its deepening concern about the water quality effects of nonpoint sources, stating that "*it is the national policy that programs for the control of nonpoint sources of pollution be developed and implemented in an expeditious manner so as to enable the goals of this Act to be met through the control of both point and nonpoint sources of pollution*." As part of the implementation of this policy, Congress added the new Section 319 and 402 to the law. One of the major features of Section 319 is its call for states to prepare, and submit to the Agency, Nonpoint Source Assessment Reports and Management Programs. A key requirement is for the states to identify those receiving waters that cannot meet the goals of the Act due to the effects of nonpoint source discharges.

Technical Assistance is an integral component of the Agency's National Nonpoint Source Program's Agenda for Action, prepared by the Nonpoint Source Task Force. This report is one of a number of specific technical assistance products being prepared by the Agency to assist state and local government officials with the planning and implementation of NPS programs. Under the new mandates of the Water Quality Act of 1987 states are given the primary responsibility for identifying receiving waters that are being impacted by urban runoff and other nonpoint sources, and for developing management programs to control these nonpoint sources so that the goals of the Act can be met.

This report is addressed to those engaged in water quality management planning and implementation, and is intended to enhance their ability to deal with such issues as establishing receiving water body priorities, determining the significance of urban runoff discharges, performing point versus nonpoint source comparisons, and determining the effect of control measures on water quality. To facilitate the utilization of the methodology presented in this report, an interactive, user-friendly computer program that performs the analysis is included as part of this report.

Users are strongly encouraged to read this document carefully. A valid technique with deceptively attractive simplicity has a greater than normal potential for misapplication and use beyond its intended capabilities. Despite its ease of use, the Agency does not believe that this method should be used by individuals who do not fully understand the screening level approach and the assumptions inherent therein, nor should it be applied to situations for which it is not technically suitable.

1.3 ORGANIZATION OF THIS REPORT

Chapter 2 of this report provides a technical overview, compares different approaches to urban runoff modeling, and outlines the rationale behind the development of the methodology and the benefits of a probabilistic approach. The theoretical basis of the methodology, including the development of equations and the assumptions made to

facilitate its use, is presented in detail in Chapter 3. Chapter 4 addresses the application of the methodology including data needs and sorces of data; computational techniques; and the interpretation of the results. Field verification of the methodology and conclusions of the study are discussed in Chapter 5. Chapter 6 presents guidance for the use of the computer program and includes example input and output files. Additional mathematical details and derivations are presented in Appendix A while Appendix B presents step-by-step calculations based on the Derived Distribution Approach. Appendix C and D include the list of variable and subroutines used by the computer program.

Chapter 2.0

TECHNICAL OVERVIEW

2.1 INTRODUCTION

This chapter presents an overview of different approaches used for urban runoff modeling and some considerations that lead to the selection of the methodology presented in this report, for estimating the impacts of stormwater discharges from urban areas on rivers and streams. Before discussing the different approaches to modeling urban runoff, however, there are some contextual issues that must be addressed. These are to define both the system being simulated and the management decision that will be based on the outputs of the chosen methodology.

2.2 THE SYSTEM BEING SIMULATED

Prior to the selection or the development of a model, it is essential that the various components of the system being simulated be identified. The model presented in this report addresses the urban runoff system. This is characterized by various components: (i) inputs to the system, (ii) the land/conveyance system, (iii) outputs from the system, (iv) the receiving water body, and (v) impact of the outputs on the receiver quality and the use. Identification of these components of the system is an essential first step in the model selection and development process.

In the present case, the inputs to the system are the meteorological conditions and in particular the precipitation rates. The land/conveyance system is comprised of the catchment, characterized by the slope and soil characteristics, type of land use and associated activities, and natural and engineered components including lakes and rivers, detention basins, highways, etc. The urban runoff system includes a number of point and nonpoint source discharges; that can be quantified as time series of water and pollutant masses. These outputs from the catchment are discharged into a receiving water body (lake, stream) for which there is some water quality concern. The water quality of the receiving water body reflects the combined characteristics of the receiver and the outputs, and impacts the beneficial uses of the receiver. A literature review indicates that each of these elements has been modeled individually as well as a coupled system.

2.3 RELEVANT MANAGEMENT DECISION

A somewhat simplistic formulation of the management decision to be addressed by the methodology presented in this report is (i) whether urban runoff is a significant contributor of pollutants in a particular river or stream, and (ii), what level of control (reduction in runoff) will be required in order to meet the specified water quality goals.

In situations where the stormwater discharges from an urban area cause an obvious negative impact on the realization of a designated beneficial use (e.g., use impairment or denial), the answer to the first decision is obvious. However, in most cases, such a determination is less obvious. In such cases, it is necessary to use a quantitative methodology, such as the one presented here. Typically, water quality criteria are used as surrogate measures of the degree to which the level of use of the receiving water body is

achieved. Thus, decisions on the impact of urban stormwater discharges upon beneficial uses can be made based upon in-stream water quality that has been estimated using a quantitative model.

It is important that the information provided by a model be appropriate to the decision to be made. Favorable consideration should be given to models that provide information for a number of different alternatives as opposed to those that are restrictive in their capabilities. For example, a model that estimates only end-of-pipe discharge information for a "design storm" is much less desirable than one that can estimate the continuum of discharges. Similarly, a model that can also estimate the effect on receiving water conditions is even more desirable.

Other important considerations are the time and spatial scales associated with the decision problem being addressed. Time scales for model outputs range from the duration of an individual event to long-term steady-state models with time scales of eternity, i.e., the entire continuum. For the decision being addressed here, models that simulate long-term conditions are particularly relevant. Similarly, spatial scales over which models operate range from micro-catchment, through catchment and watershed, to political jurisdiction (e.g., city, county, state). The methodology presented herein is appropriate for scales ranging from the size of the watershed to jurisdictional spatial scales.

2.4 CLASSIFICATION OF MODELS

There are a number of ways in which models can be classified. One is to classify them according to their formulation and use as either (i) physical/mechanistic models or (ii) decision models for either screening level anlaysis or detailed analysis. Physical models focus on relating cause and effect in terms of the physical variables of the phenomena involved and the processes being simulated. Decision models focus on relating cause and effect in terms of the decision variables of the alternatives used to control the processes being modeled.

Generally, a decision model is less appropriate than a physical model for explaining the cause and effect relationship. However, it is more appropriate for evaluating the cost-effectiveness of management options.

It must also be recognized that physical models and decision models are implemented under different conditions. Most physical models simulate details of the process, including specific water quality phenomena, under very specific prescribed conditions (e.g., a "design" storm, 7Q10). Thus, for such models the assumption is that the system behavior is known in sufficient detail though possibly for very limited conditions. On the other hand, screening level decision models simulate an aggregate of processes for a range of conditions (e.g., the entire probability distribution of flow conditions in the receiver) so that the system behavior is known in less physical detail, but for a wider range of conditions.

Thus, physical models are generally limited to the analysis of specific events, while decision models analyze a continuum of events. The trade-off in model selection is between more detailed analyses of behavior for less general conditions and less detailed analyses of behavior for more general conditions.

In order to address the decision problem, it is necessary that the level of information is most appropriate for the decision at hand. Is it preferable to have detailed information on what happens at the end-of-the-pipe (which may then be used as an acceptable surrogate for in-stream water quality) or information on resulting water quality, or even information on the level of beneficial use? As we move from input to use, we gain in detail but lose in scope. There is always a trade-off between detail and scope.

2.4.1 Approaches to Runoff Modeling

Even a cursory review of the literature indicates that a number of different modeling approaches are available (EPA 1985a and b). These modeling approaches can be grouped into the following categories: (i) Regression methods, (ii) Physically-based deterministic models, and (iii) Probabilistic techniques.

These different methods are complimentary to each other and represent varying levels of complexity and detail. The following sections present brief discussions of well-documented and tested examples of each of these three types of models and evaluate their potential use for predicting pollutant discharges from urban areas. A detailed description and evaluation of each type of model can be found in Miracle (1986).

2.4.2 Regression Models

Regression models provide a quick and simple means for estimating urban runoff quantity and quality. They are based on the assumption that explanatory variables, such as drainage area, basin imperviousness, and mean annual rainfall, can explain variation in stormwater discharge loads on a specified time scale such as long-term, annual or seasonal. The typical procedure is to perform a regression on natural or transformed data set(s) and to determine the percent of the variability (usually expressed as R^2) that can be explained by the variables used in the regression. Although causal hypotheses may be advanced to justify variable inclusion, actual physical mechanisms are not involved, and actual cause/effect relationships cannot be determined. Caution should be used when applying regression models outside of the particular study for which they were developed.

Two additional general comments should be made. First, the use of transforms can alter significantly the form of the regression equation and can introduce bias that must be properly corrected for. For example, the use of the log transform converts an additive regression model to a multiplicative one and, although the correct median is preserved, it introduces a bias in the predicted mean. The second point has to do with spurious correlations. If one regresses loads against rainfall there will be a spurious self-correlation because the load is the product of discharge quantity and concentration, and discharge quantity can be expressed as the product of rainfall and a runoff coefficient. Thus one can say, without even performing a regression, that rainfall quantity will be the single most important variable in predicting runoff loads. The recent United States Geological Society (USGS) study cited below came to this conclusion. The correct way is to perform the regression on concentrations, not loads, and thus avoid the spurious effect. Note that a correlation coefficient based upon transformed data does not accurately reflect the correlations between untransformed data.

2.4.3 Simulation Models

Numerous models have been developed to simulate stormwater runoff quantity and quality from urban and nonurban areas. However, there are relatively few models that are truly operational; that is, models that have documentation, user support, support by a government agency, and are widely used by more than just the model developers. Models meeting these requirements are SWMM (EPA's Stormwater Management Model), STORM (U.S. Corps of Engineers, Storage, Treatment, Overflow, Runoff Model, HEC (1977)), and HSPF (EPA's Hydrologic Simulation Program—Fortran, Johanson, et al. 1980). SWMM and STORM include routines to compute the effect of certain controls.

Another example is the USGS model developed for urban quality predictions by Alley and Smith (1982). However this model has not been applied as widely as SWMM, STORM, or HSPF, and its quality routines are derived basically from those in SWMM and HSPF (Huber 1985). Similarly, the Illinois State Water Survey's model ILLUDAS (Illinois Urban Drainage Simulator), the Southeast Michigan Council of Government's model RUNQUAL (Runoff and Quality Model), the Massachusetts Institute of Technology's model, MITCAT (M.I.T. Catchment Model), EPA's model HVM-QQS (HVM-Quantity-Quality Simulator) have not been used widely.

Generally, these models simulate the buildup of pollutants during dry periods, followed by washoff during storms. This procedure was first implemented in the original EPA Storm Water Management Model, SWMM (Metcalf and Eddy, et al. 1971). The formulation of the buildup and washoff equations varies from model to model, ranging from highly flexible in the SWMM to rigid in STORM. Another approach, besides modeling buildup and washoff, is the use of a rating curve methodology. This is an option in SWMM.

It is expected that complex simulation models should represent an improvement over simple regression equations for prediction of runoff quality, however, simulation models are very data intensive. These models may be viewed as "very complex regression equations" that must be calibrated in the same way that least squares is used to fit ordinary regression relationships to data. In addition to the advantage that simulation models provide a better estimate of storm event loads, these models provide a more physically realistic predictive mechanism which, when calibrated, may be altered more easily to examine the effects of changes and abatement practices. Also, simulation models can be used to estimate temporal and spatial distribution generally not possible using regression techniques. For example, continuous simulation models may be used to generate a time history of pollutant loads from which a frequency analysis may be conducted. This permits an analysis on the basis of pollutant characteristics rather than on rainfall or runoff characteristics. (The probabilistic method, described below, approaches this same end by means of a derived distribution, rather than simulation.)

2.4.4 Probabilistic Methods

Probabilistic methods are based on the realization that nonpoint sources in general, and urban runoff in particular, provide highly variable inputs (concentrations and loads) to receiving water bodies. They are generated intermittently at highly variable intervals, and when they do occur, are highly variable in both time and space. Furthermore, the impacts of nonpoint sources are strongly influenced by significant differences in important

characteristics of individual receiving water bodies and by the inherent variability of natural water systems.

For these reasons we have chosen to develop a probabilistic methodology, described in detail in subsequent chapters, for analyzing the water quality effects of urban runoff on rivers and streams. Our methodology starts with classical deterministic water quality models of pollutant fate and uses probability distributions of the relevant input parameters to estimate the probability distribution of instream concentrations. The critical situations (e.g., low flows, high loads) are all considered, but they are given their proper weight as dictated by their observed frequencies of occurrence and intercorrelations. The approach, therefore, does not depend upon the prespecification of critical conditions for the analysis; a significant advantage over the other approaches.

Probabilistic methods have the advantage of being able to consider either very long or short periods as appropriate. Sufficient data for input parameters must be available to provide reasonable confidence in the distributions developed. This is especially true if extreme events are of concern, and a high degree of confidence in their values is desired. Note that monitoring constraints are much less restrictive than those required by simulation models. Further, probabilistic data (e.g., streamflow, effluent characteristics) are much more transferable than specific sequences and sites. For example, the assembly of data to permit reliable estimates of discharge characteristics of urban runoff, CSOs (Combined Sewer Overflows), and POTW (Publicly Owned Treatment Works) discharges utilizing a specific process design will allow a rational, compatible examination of point versus nonpoint source trade-offs.

2.5 CHARACTERIZATION OF RAINFALL AND RUNOFF EVENTS

2.5.1 The Rainfall/Runoff Event

The rainfall or runoff may be viewed as a series of independent, randomly occurring events. Each event can be schematized as a uniform, rectangular hydrograph, characterized by its duration, volume, average flow rate or intensity, and the time elapsed since the last event (interevent time). The interevent time is measured between the midpoints of event. Rainfall parameters for a given site may be estimated from historical records obtained from National Oceanic and Atmospheric Administration (NOAA), e.g., by using the SYNOP program (WCC 1988).

2.5.2 Selection of Interevent Duration

The usual approach used for convenience in processing hourly rainfall data is to choose a minimum interevent time, such that rainfall values separated by less than this duration are considered part of the same storm. Values separated by times greater than or equal to this duration are considered to be independent. Several methods for choosing this duration exist (Heaney, et al. 1977), but the most common is to assume that the interevent times are exponentially distributed (Hydroscience 1979; Restrepo-Posada and Eagleson 1980). An exponential distribution is a special case of the gamma distribution when the coefficient of variation is unity. Note that the gamma distribution is typical of rainfall parameters. The SYNOP program iteratively estimates the interevent period until the coefficient of variation of time between event midpoints equals 1.0. Resulting values are usually in the range of 3 to 24 hours (Hydroscience 1979).

2.5.3 Estimation of Runoff Statistics from Rainfall Statistics

2.5.3.1 Quantity Statistics. As a first approximation the runoff coefficient method is used to calculate the mean runoff event volume and flow from the corresponding rainfall statistics.

$$VR = Rv * VP \tag{2.1}$$

where:

VR = mean runoff event volume [L^3]
Rv = average runoff coefficient [dimensionless]
VP = mean rainfall event volume [L^3]

The runoff coefficient, Rv, represents the ratio of average runoff to rainfall and neglects depression storage. This ratio varies from storm to storm, but for preliminary analyses a representative constant value is usually adequate (Nix 1982). The value of Rv can be estimated by an analysis of local rainfall/runoff data or estimated from one of several simple techniques (Hydroscience 1979).

The mean runoff flow rate is calculated as:

$$QR = Rv * QP * (DP/DR) \tag{2.2}$$

where:

QR = mean runoff event flow rate [L^3/T]
QP = mean rainfall event intensity [L^3/T]
DP = mean rainfall event duration [T]
DR = mean runoff event duration [T]

The ratio DP/DR is included to account for the fact that runoff continues after the rainfall event has subsided. A method for estimating DR using unit hydrograph analysis is presented by Hydroscience (1979).

The mean interevent time for runoff events, TR, is assumed to equal the rainfall duration, TP. The coefficients of variation for runoff event flows, CVQR, and volumes, CVVR, are also assumed to equal to their rainfall counterparts.

2.5.3.2 Quality Statistics. The mean pollutant load for all runoff events is determined by the mean pollutant concentration and the mean runoff event volume:

$$MR = CR * VR \tag{2.3}$$

where

MR = mean runoff event pollutant load [M]
CR = mean runoff event concentration [M/L^3]

This equation assumes that the pollutant event mean concentrations are independent of runoff volumes. If this is not true, an adjustment is available as discussed below. Similarly, the mean pollutant load rate can be estimated by:

$$WR = CR * QR \quad (2.4)$$

where:

WR = mean runoff event pollutant load rate [M/T]

Again, the assumption of independence between pollutant concentrations and runoff flow is made. If the independence assumptions are inadequate, the following corrections can be made (Hydroscience 1979).

$$MR = CR * VR * (1 + CVCR * CVVR * R_{cv}) \quad (2.5)$$

$$WR = CR * QR * (1 + CVQR * CVQR * R_{cq}) \quad (2.6)$$

where:

CVCR = coefficient of variation for runoff event mean pollutant concentration [dimensionless]

R_{cv} = linear correlation coefficient between pollutant concentrations and runoff volumes [dimensionless ranging from -1 to +1]

R_{cq} = linear correlation coefficient between pollutant concentrations and runoff flow rates [dimensionless]

A positive value of R_{cv} or R_{cq} indicates that higher flows or volumes produce higher concentrations. A negative value indicates that the dilution effect of large runoff events is dominant.

Finally, it is assumed that all the variation in loads is due to variation in runoff volumes. Thus,

$$CVMR = CVVR \quad (2.7)$$

and

$$CVWR = CVQR \quad (2.8)$$

where

CVMR = coefficient of variation of runoff pollutant loads [dimensionless]
CVWR = coefficient of variation of runoff pollutant load rates [dimensionless]
CVVR = coefficient of variation of runoff volume [dimensionless]
CVQR = coefficient of variation of runoff flow [dimensionless]

The probability distribution of loads and concentrations has been assumed to be gamma for application to certain storage-treatment analysis (Di Toro and Small 1979; Small and Di Toro 1979) but is usually taken to be lognormal for application to stream water quality impacts (Di Toro 1984).

2.6 SUMMARY

Regression equations, simulation models and the probabilistic method are analytical tools for predicting pollutant loads from urban runoff. The primary advantage of

the regression and probabilistic methods is that both allow a relatively quick, simple, and inexpensive screening of stormwater problems. This is particularly useful in the early stages of the planning process (Hydroscience 1979).

Each method has its own particular advantages and disadvantages. The regression methods are simple and include explanatory variables such as rainfall characteristics, traffic counts, population density, etc. However, this method can yield only the mean for a given set of inputs and their range of validity is generally limited. Implementation of simulation models requires considerable resources, especially in terms of calibration and verification data requirements, and are not recommended for screening type analysis.

The probabilistic method requires rainfall statistics (readily available) and known mean concentrations of pollutants in the runoff (somewhat harder to obtain) as input variables (plus streamflow statistics, if receiving water effects are to be investigated). The output is generally a frequency distribution of a water quality parameter and can be used to estimate return periods, i.e., the frequency that a given concentration level be equalled or exceeded. The method includes some approximations if an analytical solutionis desired, however, a few of which may be important in limited individual cases (Roesner and Dendrou 1985).

Based on the foregoing, the probabilistic approach has been selected for screening level analysis. Its success in EPA's NURP and NOAA's National Coastal Discharge Inventory Program, as well as the Federal Highway Administration (FHWA) Runoff Program, and its increasing use by state water quality planners and other professionals at the local level further encourage its selection. Also, as noted earlier, it has been exhaustively reviewed by EPA's Science Advisory Board, who found the method to be technically sound.

Probabilistic methodology provides a method to deal with the inherent variablity of the stormwater runoff process and its impacts on the receiving water. It also emphasizes the impossibility of reliably predicting pollutant loads based on an indivdual storm event explicitly quantifies the uncertainty involved in the estimation of individual site pollutant runoff. These characteristics of the probabilistic approach provide a basis for evaluating the confidence level of the decision involved.

CHAPTER 3.0

THEORETICAL BASIS OF MODEL

3.1 INTRODUCTION

Urban runoff, stream hydrology, and related water quality processes are inherently random phenomena. Thus, it is appropriate to analyze water quality effects using a probabilistic framework. The three approaches, as described below use the probability distribution of model inputs (e.g., stream discharge, runoff flow, and concentrations) to estimate the probability distribution of the resulting concentrations in the stream. Thus, if the level of concentration that impairs the use of the receiving water is known, this method can be used to estimate the frequency with which that use will be impaired.

The probabilistic water quality analysis starts with a deterministic mass balance model, i.e., a mathematical representation of the mixing of surface runoff with the river or stream and uses three different approaches. These include (i) the Analytical Derived Distribution Approach, (ii) the Direct Numerical Solution Approach (Gaussian Quadrature Method), and (iii) the Monte Carlo Simulation approach to compute the probability distribution of the stream concentrations. Critical situations (e.g., low flows, high loads, etc.) are considered, and given their proper weight as dictated by their observed frequency of occurrence.

3.2 MODEL FRAMEWORK AND STATISTICAL CONCEPTS

The conceptual formulation consists of a spatially lumped, event-averaged model. All concentrations and flows are event mean values. Although urban runoff discharges usually enter the receiving water at several locations, the model considers the aggregate of flows and assumes complete mixing in the stream. Figure 3-1 shows an idealized representation of urban runoff discharges entering the river. Downstream concentrations are estimated using mass balance or a dilution model:

$$C_0 = \frac{Q_S}{Q_S+Q_R} C_S + \frac{Q_R}{Q_S+Q_R} C_R \tag{3.1}$$

where:

C_0 = downstream concentration in the river [mg/l]
C_R = pollutant concentration in the runoff [mg/l]
C_S = upstream concentration in the river [mg/l]
Q_R = runoff flow rate [m^3/s]
Q_S = upstream flow rate in the river [m^3/s]

The objective of the methodology presented here is to estimate the probability distribution of C_0, for specified probability distributions of the input parameters Q_S, Q_R, C_R, and C_S as shown in Figure 3-1.

The three probabilistic methodologies presented below, consider C_S, Q_S, C_R, and Q_R as random variables, each specified as a probability distribution. As described in the

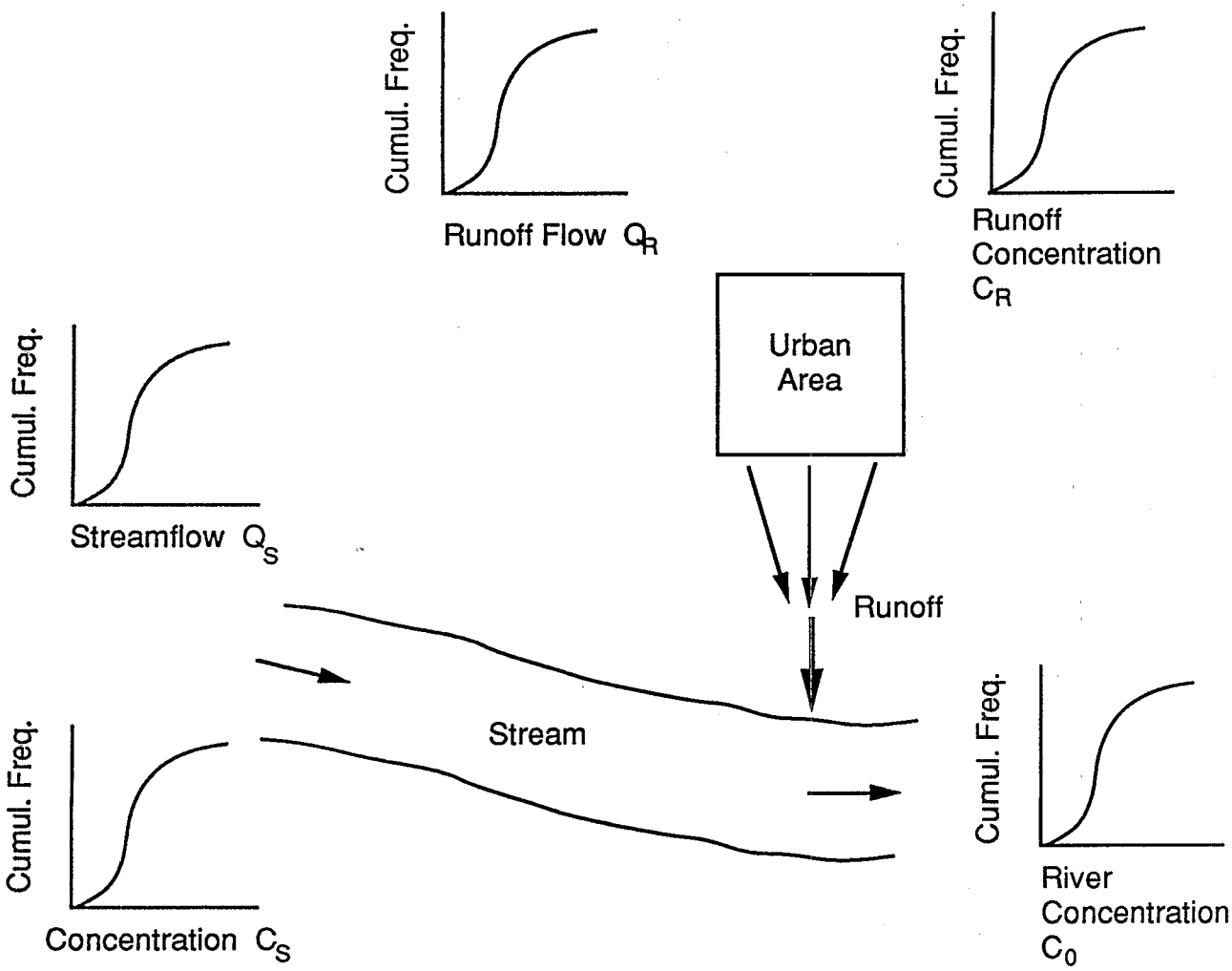


FIGURE 3-1

SCHEMATIC REPRESENTATION OF WATER QUALITY EFFECTS OF URBAN RUNOFF ON RIVERS AND STREAMS

NURP final report (WCC 1983), Medina (1982), and numerous other reports related to the distribution underlying variables that describe natural system, these variables can be characterized as lognormal distributions. The first two methods presented below the Analytical Derived Distribution Approach and the Direct Numerical Solution Approach are based on this assumption.

The use of probability distributions other than log normal has been proposed in the literature to represent highly variable water quality parameters. For example, Rossman (1983) presented an approach to estimate water quality criteria violations due to combined sewer overflows using Pearson percentiles. If site-specific data suggest that distributions other than log normal are more appropriate, the third approach, a Monte Carlo simulation technique may be used.

In Section 3.3 and 3.4, the Analytical Derived Distribution Approach for computing the distribution of C_0 is presented. Although this method gives approximate results, it is amenable to hand calculation and hence very useful. A numerical Gaussian Quadrature method is presented in Sections 3.5 and 3.6. Sections 3.7 and 3.8 present the Monte Carlo Simulation Approach. Finally, a useful transformation of the probabilistic results into a more meaningful, time-based recurrence interval is described in Section 3.9.

3.3 THE ANALYTICALLY DERIVED DISTRIBUTION APPROACH

The analytically derived distribution approach is an analytical procedure to estimate the exceedance probability of the downstream concentration and is based on the assumption that this concentration is lognormally distributed. It involves the computation of the first two moments (i.e., the mean and variance) of the downstream concentration.

The lognormality assumption of downstream concentration follows from the following observations. The dilution model (i.e., mass balance equation) can be written in the form:

$$C_0 = C_R \phi + C_S (1 - \phi) \tag{3.2}$$

$$\phi = \frac{Q_R}{Q_R + Q_S} \tag{3.3}$$

where

ϕ = the runoff flow fraction or dilution [dimensionless].

If ϕ and 1-ϕ were lognormal, then, since C_R and C_S are assumed to be lognormal, the products $C_R \phi$ and $C_S (1 - \phi)$would also be lognormal. Further it is assumed that the sums of lognormal random variables are also approximately lognormal (Janos 1970). These observations suggest that if ϕ is approximately lognormal, then C_0 should be approximately lognormal, and the mean and variance suffice to determine its distribution. This approach is approximate because, in fact, ϕ has an upper bound of 1 and is thus not truly lognormal. While ϕ may be appropriately considered lognormal in many cases, this assumption may introduce unreasonable errors in other cases, and one of the approaches presented in Sections 3.5 through 3.9 must be used.

With the above assumptions, the arithmetic mean and variance of C_0 are given by:

$$\mu(C_0) = \mu(C_R)\,\mu(\phi) + \mu(C_S)\,[1 - \mu(\phi)] \tag{3.4}$$

$$\sigma^2(C_0) = \sigma^2(\phi)[\mu(C_R) - \mu(C_S)]^2 + \sigma^2(C_R)\,[\sigma^2(\phi) + \mu^2(\phi)]$$

$$+\sigma^2(C_S)\,[\sigma^2(\phi) + (1-\mu(\phi))^2] \tag{3.5}$$

where

$\mu(x)$ = arithmetic mean of the variable x
$\sigma(x)$= standard deviation of the variable x

Note the above further assumes that both concentrations, C_S and C_R, are uncorrelated with respect to their respective flows, Q_S and Q_R, and, hence with the runoff fraction, ϕ.

Assuming that C_0 is lognormal, the following relationships between the arithmetic and the logarithmic (normal) moments (Aitchison et al. 1957) can be used to derive the logarithmic mean and variance of C_0:

$$\mu_{ln}(C_0) = \ln\,[\mu(C_0)/(1 + \nu^2(C_0))^{0.5}] \tag{3.6}$$

$$\sigma^2_{ln}(C_0) = \ln\,[1 + \nu^2(C_0)] \tag{3.7}$$

where

$\mu_{ln}(C_0)$ = mean of C_0 in the logarithmic (normal) space
$\sigma_{ln}(C_0)$ = standard deviation of C_0 in the logarithmic (normal) space
$\nu(C_0)$ = coefficient of variation of C_0 in the arithmetic space and is given by

$$\nu(C_0) = \frac{\sigma(C_0)}{\mu(C_0)} \tag{3.8}$$

Further, the quantiles of C_0 can be estimated using:

$$C_{0\alpha} = \exp\,[\mu_{ln}(C_0) + z_\alpha \sigma_{ln}(C_0)] \tag{3.9}$$

where

$C_{0\alpha}$ = the downstream concentration exceeded with a probability $(1 - \alpha)$ [mg/l]
z_α = the standard normal α quantile [dimensionless]

In order to use Equations 3.4 and 3.5, it is necessary to compute the moments (i.e., mean and variance) of ϕ. Warn and Brew (1980) suggest a numerical method. However, for screening level analysis, it is useful to find approximate expressions for these moments and to use the numerical techniques for the evaluation of the exact distribution of C_0. The runoff fraction Equation 3.3 can be expressed as:

$$\phi = \frac{1}{1+D} \tag{3.10}$$

where

$$D = \frac{Q_S}{Q_R} \tag{3.11}$$

If both Q_S and Q_R are lognormal, D would be lognormal. Thus the logarithmic (normal) mean and standard deviation of D can be estimated using:

$$\mu_{ln}(D) = \mu_{ln}(Q_S) - \mu_{ln}(Q_R) \tag{3.12}$$

$$\sigma^2_{ln}(D) = \sigma^2_{ln}(Q_S) + \sigma^2_{ln}(Q_R) - 2\sigma_{ln}(Q_S)\sigma_{ln}(Q_R)\rho_{ln}(Q_S,Q_R) \tag{3.13}$$

where $\rho_{ln}(Q_S,Q_R)$ is the cross-correlation coefficient between $\ln(Q_S)$ and $\ln(Q_R)$. Using Equation 3.10 the α-quantile of ϕ can be computed (refer to Appendix A) using:

$$\phi_\alpha = \frac{1}{1+\exp[\mu_{ln}(D) - z_\alpha\sigma_{ln}(D)]} \tag{3.14}$$

where

$$P_r(\phi<\phi_\alpha) = \alpha \tag{3.15}$$

Since ϕ is assumed to be lognormally distributed, $\ln(\phi)$ is normally distributed with mean of $\mu_{ln}(\phi)$ and standard deviation of $\sigma_{ln}(\phi)$. Using this assumption, the α_1 and α_2 quantiles of ϕ can be estimated using:

$$\ln(\phi_{\alpha_1}) = \mu_{ln}(\phi) + z_{\alpha_1}\sigma_{ln}(\phi) \tag{3.16}$$

$$\ln(\phi_{\alpha_2}) = \mu_{ln}(\phi) + z_{\alpha_2}\,\sigma_{ln}(\phi) \tag{3.17}$$

Equations 3.16 and 3.17 imply:

$$\sigma_{ln}(\phi) = \frac{[\ln(\phi_{\alpha_1}) - \ln(\phi_{\alpha_2})]}{z_{\alpha_1} - z_{\alpha_2}} \tag{3.18}$$

$$\mu_{ln}(\phi) = \frac{[z_{\alpha_2}\ln(\phi_{\alpha_1}) - z_{\alpha_1}\ln(\phi_{\alpha_2})]}{z_{\alpha_2} - z_{\alpha_1}} \tag{3.19}$$

In the original development (Di Toro 1984), z_{α_1} was set equal to 1.645 and z_{α_2} was set equal to -1.645 to force agreement at the 5% and 95% quantiles i.e. by estimating the 5th and 95th percentiles of ø using Equation 3.14. In many water quality impact evaluations, obtaining a close approximation of the stream concentrations with 3-year recurrence intervals (per EPA toxic criteria) is of more practical interest than the concentration levels that occur at a weekly or monthly frequency. Results of the sensitivity analyses, discussed in Chapter 5, indicate that estimating the mean and standard deviation of ø using two values in the tail (e.g., exceedance percentiles of 90 and 99.6 percent) results in better estimates of stream concentrations in the 1 to 5 year range of recurrence intervals, for a range of values of the input parameters.

Equations 3.18 and 3.19 can then be used to estimate the logarithmic mean and standard deviation of the lognormal approximation of ø. The arithmetic mean and standard deviation of ø (required for Equations 3.4 and 3.5) can be computed as:

$$\mu(\phi) = \exp[\mu_{ln}(\phi) + 0.5\,\sigma^2_{ln}(\phi)] \tag{3.20}$$

$$\sigma(\phi) = \{\exp[\sigma^2_{ln}(\phi)]-1\}^{0.5}\mu(\phi) \tag{3.21}$$

Thus, the steps required to compute the exceedance probabilities of C_0 are:

(i) Analyze the streamflow and the runoff data to estimate the mean and standard deviation of the lognormal distributions.
(ii) Estimate the first two moments of D using Equations 3.12 and 3.13.
(iii) Estimate the values of ø for the selected percentiles using Equation 3.14.
(iv) Estimate the log mean and log standard deviation of ø using Equations 3.18 and 3.19 and the arithmetic mean and standard deviation using Equations 3.20 and 3.21 respectively.
(v) Estimate the arithmetic mean and standard deviation of C_0 using Equations 3.4 and 3.5.
(vi) Estimate the desired percentile of C_0 using Equation 3.9.
(vii) Estimate the mean recurrence interval using Equation 3.43 described in Section 3.7.

In Appendix B, an example calculation is given. The accuracy of this approximate procedure is examined in Chapter 5.

3.4 KEY ASSUMPTIONS FOR THE ANALYTICALLY DERIVED DISTRIBUTION APPROACH

The following assumptions are inherent in the probabilistic methodology presented above.

- All flows (Q_R,Q_S) and pollutant concentrations (C_R,C_S) are lognormally distributed.
- The runoff and upstream concentration (C_R and C_S) are assumed to be uncorrelated with the runoff flow and streamflow (Q_R and Q_S) respectively, as well as with the parameters ø and (1-ø).
- The sum of C_Rø and C_S(1-ø); i.e., C_0 is assumed to be lognormally distributed.
- The dilution parameter ø (refer to Equation 3.3) and (1-ø) are both assumed to be lognormally distributed.

In addition, the mathematical model (Equation 3.1) uses the following assumptions:

- The stream is assumed to be fully mixed both upgradient and downgradient of the urban runoff discharges.
- The runoff from the urban area is aggregated into a single event mean flow and concentration value.

3.5 THE DIRECT NUMERICAL SOLUTION

The probability of downstream concentration, C_0 , exceeding a particular value of interest, C^*_0, can be expressed as a multiple integral of the joint probability density function over the values of flows and concentrations for which $C_0 > C^*_0$. This requires an integration for each of the basic model variables C_R, Q_R, C_S, Q_S. One integral can be eliminated by using D as defined by Equation 3.11. Thus:

$$P_r\{C_0 > C_0^*\} = P_r\{\frac{C_R}{1+D} + \frac{C_S D}{1+D} > C_0^*\} \tag{3.22}$$

Since Q_S and Q_R are both lognormal, so also is D. For a fixed C_0* the equality:

$$C_0^* = \frac{C_R}{1+D} + \frac{C_S D}{1+D} \tag{3.23}$$

defines a surface in C_R, C_S, and D space; and the required probability is the integral of the joint probability density function over this surface. Specifically, for a fixed D:

$$C_R + C_S D = C_0^*(1+D) \tag{3.24}$$

is a straight line in the C_R, C_S plane as shown in Figure 3-2. Note that the inequality in equation 3-22 is valid in the region to the right of the straight line given by equation 3-24. This region can be divided into two parts Region A and Region B defined by:

<u>Region A</u>

$$0 \le C_S \le C_0^* \left(1 + \frac{1}{D}\right)$$
$$C_0^* (1 + D) - C_S D \le C_R \le \infty$$

<u>Region B</u>

$$C_0^* \left(1 + \frac{1}{D}\right) \le C_S \le \infty$$
$$0 \le C_R \le \infty$$

The above is valid for the entire range of values of D; i.e., D can vary from $[0,\infty]$. In order to calculate the exceedance probability shown in equation 3-22, it is necessary to integrate over the regions A and B. Thus:

$$P_r\{C_0 > C_0^*\} = \underbrace{\int_{D=0}^{\infty}\int_{0}^{C_{S1}}\int_{C_{R1}}^{\infty} f(C_R,C_S,D)\, dC_R\, dC_S\, dD}_{\text{Region A}} + \underbrace{\int_{0}^{\infty}\int_{C_{S1}}^{\infty}\int_{0}^{\infty} f(C_R,C_S,D)\, dC_R\, dC_S\, dD}_{\text{Region B}} \quad (3.25)$$

where:

$$C_{R1} = C_0^* (1+D) - C_S D \quad (3.26)$$

$$C_{S1} = C_0^* \left(1+\frac{1}{D}\right) \quad (3.27)$$

and $f(C_R, C_S, D)$ is the joint probability density function for C_R, C_S and D.

Since C_R, C_S and D are assumed to be lognormal, a change of variables can be made and the joint probability density function in Equation 3.25 can be expressed as a trivariate gaussian. A convenient method of expressing this density is the product of the marginal and conditional densities, (Anderson 1958), each of which is univariate gaussian. Thus:

$$f(c_R,c_S,d) = f(d)\, f(c_S \mid d)\, f(c_R \mid c_S,d) \quad (3.28)$$

where:

$$c_R = \ln C_R \quad (3.29)$$

$$c_S = \ln C_S \quad (3.30)$$

$$d = \ln D \quad (3.31)$$

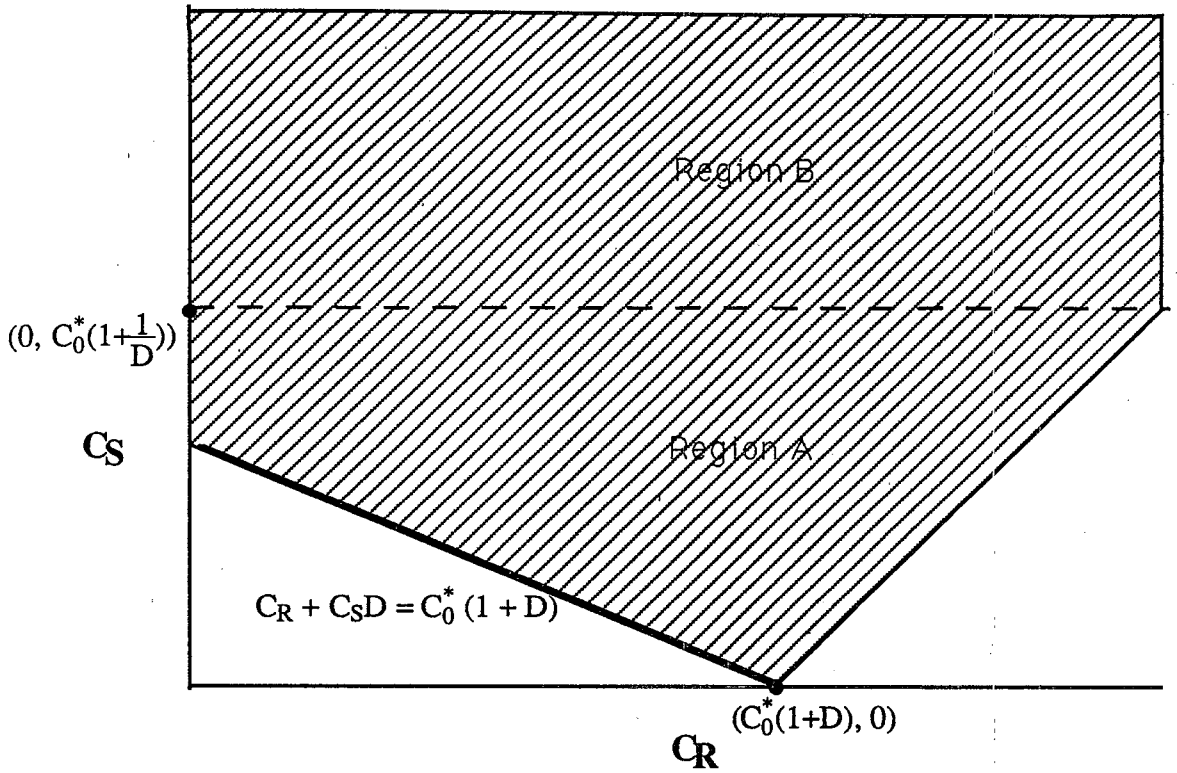


FIGURE 3-2

REGIONS OF INTEGRATION FOR THE JOINT PROBABILITY DENSITY FUNCTION

$$f(d) = N[d;\ \mu(d),\ \sigma(d)] \tag{3.32}$$

$$f(c_s|d) = N[c_s;\ \mu(c_s|d),\ \sigma(c_s|d)] \tag{3.33}$$

$$f(c_R|c_s,d) = N[c_R;\ \mu(c_R|c_s,d),\ \sigma(c_R|c_s,d)] \tag{3.34}$$

Note Equations 3.32, 3.33 and 3.34 are the univariate gaussian densities with the appropriate conditional logarithmic (normal) means and standard deviations. Equation 3.25 can now be written as:

$$P_r\{C_0>C_0^*\} = \int_{-\infty}^{\infty} f(d)\ \{\int_{-\infty}^{\ln C_{S1}} f(c_S|d) \int_{\ln C_{R1}}^{\infty} f(c_R|c_S,d)\ dc_R\ dc_S dd + \int_{-\infty}^{\infty} f(d) \int_{\ln C_{S1}}^{\infty} f(c_S|d) \int_{-\infty}^{\infty} f(c_R|c_S,d)\ dc_R dc_S\}\ dd \tag{3.35}$$

Equation 3.35 can be simplified using the following relationships:

$$\int_{-\infty}^{\infty} f(c_r|c_S,\ d)\ dc_r = 1$$

$$\int_{\ln C_{S1}}^{\infty} f(c_S|d) = P_r\ \{C_S > C_{S1}\} = Q_S^*$$

$$\int_{\ln C_{R1}}^{\infty} f(c_r|c_S,d) = P_r\ \{C_r > C_{R1}\} = Q_R^*$$

This results in the following equation:

$$P_r\{C_0>C_0^*\} = \int_{-\infty}^{\infty} f(d) \int_{-\infty}^{\ln C_{S1}} f(c_S|d)\ Q_R^*\ dc_S dd + \int_{-\infty}^{\infty} f(d) Q_S^*\ dd \tag{3.36}$$

Q_R^* and Q_S^* are evaluated using the following equations:

$$Q_R^* = Q^*\ [\frac{\ln(C_{R1}) - \mu(c_R|c_s,\ d)}{\sigma(c_R|\ c_S,\ d)}] \tag{3.37}$$

$$Q_S^* = Q^*\ [\frac{\ln(C_{S1}) - \mu(c_S|\ d)}{\sigma(\ c_S|\ d)}] \tag{3.38}$$

where $Q^*(z^*) = P_r(z>z^*)$, the probability that a zero mean, unit variance gaussian random variable exceeds z^*. This function can be computed using approximate expressions (Abramowitz and Stegun 1954).

The remaining integrals can be evaluated numerically after a transformation of the variable as discussed in Appendix A. The integral is solved using standard gaussian quadrature formulae.

3.6 KEY ASSUMPTIONS IN THE DIRECT NUMERICAL SOLUTION

The numerical methodology presented above is based on a number of assumptions. These assumptions are listed below:

- The stream is assumed to be fully mixed both upgradient and downgradient of the urban runoff discharges.
- The runoff from the urban area is aggregated into a single event mean flow and concentration value.
- All flows (Q_R,Q_S) and pollutant concentrations (C_R,C_S) are lognormally distributed.

3.7 THE MONTE CARLO SIMULATION APPROACH

A commonly used approach for analyzing the uncertainty in model outputs due to uncertainty in the model inputs is the Monte Carlo simulation approach. This approach is conceptually simple, general in its applicability and asymptotically exact, but not suitable for computationally intensive models. The application of this method to the dilution model is described below.

Given a set of deterministic values for each of the input parameters, QS, QR, CS and CR, the dilution model Equation 3.1 computes the downstream concentration Co Application of the Monte Carlo simulation procedure requires that at least one of the input variables, QS, QR, CS and CR, be uncertain and the uncertainty represented by a cumulative probability distribution. The method involves the repeated generation of pseudo-random values of the uncertain input variable(s) (drawn from the specified distribution and within the range of any imposed bounds) and the application of the model using these values to generate a series of model results, i.e., values of Co. These results are then statistically analyzed to yield the cumulative probability distribution of the model response. Thus, the various steps for the application of the Monte Carlo simulation technique involve:

i) Selection of representative cumulative probability distribution functions for the model input parameters.

ii) Generation of pseudo-random numbers from the distributions selected in (i). These values represent a possible set of values for the input variables.

iii) Application of the dilution model to compute the downgradient concentrations.

iv) Repeated application of steps (ii) and (iii).

v) Presentation of the series of downstream concentrations generated in step (iii) as a cumulative probability distribution function (CDF).

The random number generators included in the software allows the user to select one of the following distributions for the input parameters:

i) Normal
ii) Lognormal
iii) Uniform
iv) Exponential
v) Empirical

The first two distributions require the user to specifiy the mean and the standard deviation. The third requires minimum and maximum values. The fourth distribution requires only one parameter - the mean of the distribution. Finally for the empirical distribution, the user is required to input the coordinates of the cumulative probability distribution function which is subsequently treated as a piece-wise linear curve.

3.8 KEY ASSUMPTIONS OF THE MONTE CARLO SIMULATION APPROACH

The Monte Carlo simulation method does not use any assumptions other than those included in the mathematical model (Equation 3.1). These are:

- The stream is assumed to be fully mixed both upgradient and downgradient of the urban runoff discharges.
- The runoff from urban discharges is aggregated into a single event mean flow and concentration value.

3.9 RECURRENCE INTERVAL CALCULATION

The fundamental result of the statistical analysis described above is the derived cumulative probability distribution of event mean concentrations downstream of the urban area. This distribution is given by:

$$C_{0\alpha} = \exp[\mu_{ln}(C_0) + z_\alpha \sigma_{ln}(C_0)] \tag{3.39}$$

where

$C_{0\alpha}$ = the concentration that is exceeded with probability $1-\alpha$
z_α = the standard normal quantile.

One way of interpreting the probability corresponding to a given concentration level (i.e., $P_r\{C_0 \geq C^*_0\}$) is the long term average fraction of events that a stream event mean concentration equals or exceeds the specified level. For example, a probability of 0.10 would specify that, on average, one in ten events have a downstream event mean concentration equal to or greater than the specified value.

For purposes of evaluation and interpretation, it is useful to transform the basic probability statistic into a more meaningful or intuitive form. By combining the percent of storms that cause various concentrations to be exceeded with the average number of storms per year, a time-based recurrence interval may be established as described below.

Reccurrence is based (generally) on the marginal distribution of random variables. If $P_r\{C_0 \geq C^*_0\}$ is the probability of a value of magnitude C^*_0 being equaled or exceeded in a given time period, then the mean recurrence interval (MRI) defined as the reciprocal of probability is the average number of time periods between exceedances.

Thus annual recurrence is the reciprocal of the probability that an event concentration of a given magnitude, C^*_0, is equaled or exceeded in a year. The statement of the problem is:

$$P_r\{C_0^{max} \geq C_0^*\} = P_r\{\max(C_{0,1}, C_{0,2}, \ldots, C_{0,N}) \geq C_0^*\} \tag{3.40}$$

where

C_0^{max} = the maximum event concentration in a year [mg/l]
N = the number of events in a year.

Assuming that event concentrations are independent and identically distributed with a known distribution, such as lognormal, the above equation becomes:

$$P_r\{C_0^{max} \geq C_0^*\} = 1 - P_r\{C_0 \geq C_0^*\}^N \tag{3.41}$$

Or the mean recurrence interval in years is:

$$MRI = \frac{1}{1 - P_R\{C_0 \geq C_0^*\}^N} \tag{3.42}$$

A first order approximation to this is given by:

$$MRI = \frac{1}{P_R\{C_0 \geq C_0^*\} N} \tag{3.43}$$

The mean recurrence interval in years as defined by the above equation is a convenient and meaningful way to interpret the basic probability results. In Chapter 4 the use and interpretation of recurrence interval is described and example results are given.

CHAPTER 4.0

APPLICATION OF METHODOLOGY

4.1 INTRODUCTION

This chapter presents an example application of the three probabilistic methodologies described in the previous chapter. A step-by-step description of the Analytically Derived Distributions Approach that is amenable to hand calculations is presented. The results are compared with the results obtained using the Direct Numerical Method and the Monte Carlo Simulation Technique. Since these methods are not conducive to hand calculations, a Fortran 77 PC-based computer program described in Chapter 6 was used. A comparison of the results obtained using the three methods is presented in Chapter 5.

4.2 DATA NEEDS AND SOURCES

The input data consist of the parameters of the probability distributions for streamflow and concentration of the pollutant to be analyzed, at a location immediately upstream of the urban area. Similar data are required for the urban runoff discharges, in this case based on storm event averages for runoff flow and concentration. If computations are performed with upstream concentrations of zero, the output would reflect the downstream concentrations due to urban runoff discharges.

Streamflow characteristics may be estimated by analysis of USGS daily streamflow records for an appropriate flow gaging station in the study area. Alternatively, estimates for typical regional values, suitable for preliminary estimates, can be derived from information presented in the final NURP report (WCC 1983).

Statistical parameters for upstream concentration may be computed from local monitoring data or STORET data, if the available period of record is sufficiently long.

Statistical parameters for pollutant concentrations in urban runoff may be computed by: (i) using data obtained from local monitoring programs, (ii) the information presented in the final NURP report, or (iii) by a combination of the two sources of data.

Statistical parameters for the urban runoff flows may be computed based on analyzing data from a local monitoring program. However, if such information is not available, an alternative is to estimate a runoff coefficient for the urban area being examined. This would then be combined with the long-term rainfall record for the local area to derive estimates of long-term statistics of runoff flows. Details of this method are presented in WCC (1983).

For the Derived Distribution Approach and the Direct Numerical Method, computations are performed using log transforms of the source data; however, results, evaluations, and interpretations are most appropriately presented as the relevant arithmetic values. Note that there are specific relationships between arithmetic and logarithmic parameters, so that transpositions to the desired forms can be made quite readily. These relationships are presented in Table 4-1. Note that for the two-parameter, lognormal distributions used in the methodology, specifying values for any two of the arithmetic or logarithmic parameters defines all others.

TABLE 4-1

INPUT DATA REQUIREMENTS AND RELATIONSHIPS
FOR LOGNORMAL DISTRIBUTIONS

Input Parameter		Arithmetic: Median $\tilde{x}$	Arithmetic: Mean $\mu_{(x)}$	Arithmetic: Std. Dev. $\sigma_{(x)}$	Arithmetic: Coef. Var. $\nu_{(x)}$	Logarithmic: Log Mean $\mu_{\ln(x)}$	Logarithmic: Log S.D. $\sigma_{\ln(x)}$
Streamflow	Q_S	$\widetilde{Q_S}$	$\mu(Q_S)$	$\sigma(Q_S)$	$\nu(Q_S)$	$\mu_{\ln}(Q_S)$	$\sigma_{\ln}(Q_S)$
Stream Concentration	C_S	$\widetilde{C_S}$	$\mu(C_S)$	$\sigma(C_S)$	$\nu(C_S)$	$\mu_{\ln}(C_S)$	$\sigma_{\ln}(C_S)$
Runoff Flow	Q_R	$\widetilde{Q_R}$	$\mu(Q_R)$	$\sigma(Q_R)$	$\nu(Q_R)$	$\mu_{\ln}(Q_R)$	$\sigma_{\ln}(Q_R)$
Runoff Concentration	C_R	$\widetilde{C_R}$	$\mu(C_R)$	$\sigma(C_R)$	$\nu(C_R)$	$\mu_{\ln}(C_R)$	$\sigma_{\ln}(C_R)$

RELATIONSHIPS BETWEEN STATISTICAL PROPERTIES
ARITHMETIC AND LOGNORMAL TRANSFORMS

DEFINITION OF TERMS

Arithmetic		Logarithmic
x	Random Variable	$\ln x$
μ_x	Mean	$\mu_{\ln}(x)$
σ_x^2	Variance	$\sigma_{\ln}^2(x)$
σ_x	Standard Deviation	$\sigma_{\ln}(x)$
ν_x	Coefficient of Variation	(not used)
$\tilde{x}$	Median	(not used)

$$\mu_x = e^{\mu_{\ln}(x) - 1/2\,\sigma_{\ln}^2(x)}$$

$$\mu_x = \tilde{x} * \sqrt{1 + \nu_x^2}$$

$$\tilde{x} = e^{\mu_{\ln}(x)}$$

$$\nu_x = \sqrt{e^{\sigma_{\ln}^2(x)} - 1}$$

$$\sigma_x = \mu_x \nu_x$$

$$\sigma_{\ln x} = \sqrt{\ln(1 + \nu_x^2)}$$

$$\mu_{\ln x} = \ln(\tilde{x})$$

$$\mu_{\ln x} = \ln\left(\mu_x / e^{1/2\,\sigma_{\ln}^2(x)}\right)$$

$$\mu_{\ln x} = \ln\left(\mu_x / \sqrt{1 + \nu_x^2}\right)$$

When the underlying population of a variable is presumed to be lognormally distributed, the appropriate way to analyze data drawn from a limited sample of the overall population is to convert observed values to their corresponding logarithmic values, and to perform the standard statistical analyses (determination of mean and standard deviation) on the log transforms of the data. Arithmetic statistical parameters can then be computed from the log-mean and log-standard deviation using the relationship shown in Table 4-1. The presumption is that the best estimator of the value of interest (e.g., the mean) is provided by this procedure. It should be recognized that estimates of the mean determined this way will differ from values determined from a simple averaging of the observed data, especially where sample size is small. For lognormal distributions, estimates of the mean using the two approaches will converge to a common value as the number of observations increase.

4.3 COMPUTATION PROCEDURE USING THE ANALYTICALLY DERIVED DISTRIBUTION APPROACH

The basic approach is summarized on Figure 4-1 and consists of the following four steps:

Step 1: Compute Statistical Parameters of Inputs Assuming Lognormal Distribution

Analyze local data or use NURP summaries to define the parameters of the lognormal distributions for each of the four input parameters. Published summary characteristics are usually in terms of an arithmetic mean (or median) and standard deviation (or coefficient of variation). When the input parameters are estimated from analysis of site-specific data, the log transforms of individual observations are used to estimate the log-mean and log-standard deviation. In either case, the approximate formulae listed in Table 4-1 may be used to compute the remaining statistical parameters.

Step 2: Compute the Parameters of the Flow Ratio Q_S/Q_R and the Dilution Factor

Estimate the statistical properties of the dilution factor (ø) defined by Equation 3.3. For this it is necessary to compute the statistical properties of the flow ratio (parameter D defined by Equation 3.11) using Equations 3-12 and 3-13. Equation 3.14 is then used to estimate the values of (ø) for the selected percentiles. The values of z_α in Equation 3.14 are obtained from a Standard Normal Probability table for the corresponding value of α. A few values of α and z_α are given below:

α	$1 - \alpha$	z_α
95 %	5 %	$z_{.95} = 1.645$
5 %	95 %	$z_{.05} = -1.645$
50 %	50 %	$z_{.50} = 0$
84.13 %	15.87 %	$z_{.84} = 1.0$
15.87 %	84.13 %	$z_{.16} = -1.0$
90.0 %	10.0 %	$z_{.90} = 1.282$
99.6 %	0.4 %	$z_{.996} = 2.652$

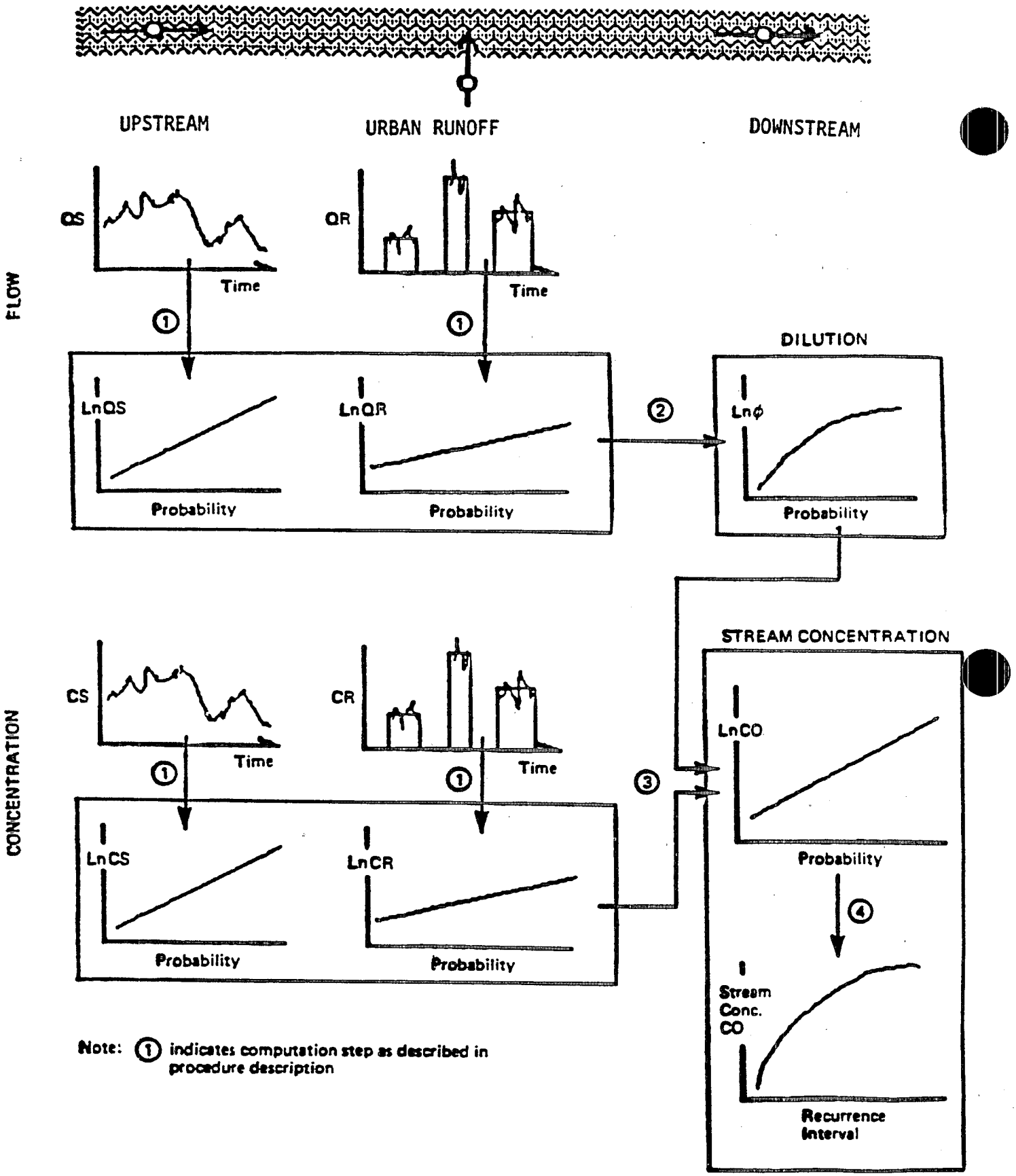


FIGURE 4-1

SCHEMATIC OUTLINE OF PROBABILISTIC ANALYSIS METHOD

These values are used to estimate the log-mean and log-standard deviation of dilution (ø) of the lognormal approximation of (ø), using Equations 3.18 and 3.19 . The arithmetic transforms are then computed using appropriate equations presented in Table 4-1. Specifically, the arithmetic mean ($\mu(ø)$) and standard deviation ($\sigma(ø)$) (Equations 3.20 and 3.21) are required in Step 3.

As discussed in Section 3.3, the probability distribution of the dilution function (ø) is not truly lognormal since it has an upper bound of 1 and a lower bound of 0. To ensure that the lognormal approximation for this distribution yields the greatest accuracy the estimates of instream concentration values in the tail of the distribution are selected. For this example, exceedance percentiles ($1 - \alpha$) of 10 and 0.4 percent are selected. The lower, more frequently occurring stream concentrations will tend to be overestimated (perhaps appreciably) by this choice. However, these values have little practical significance. The sensitivity of projections to this approximation is examined further in Chapter 5.

Step 3: Compute the Statistical Parameters of the In-Stream Concentration

The probability distribution of in-stream concentration is computed from the statistical parameters for flow, concentration, and dilution estimated in the previous steps, using Equations 3.4 and 3.5. The coefficient of variation is calculated using Equation 3.8. These statistics are used to derive the log-transformed statistical parameters using the information presented in Table 4-1.

Step 4: Prepare Probability Distributions of Downgradient Concentrations and the Recurrence Intervals

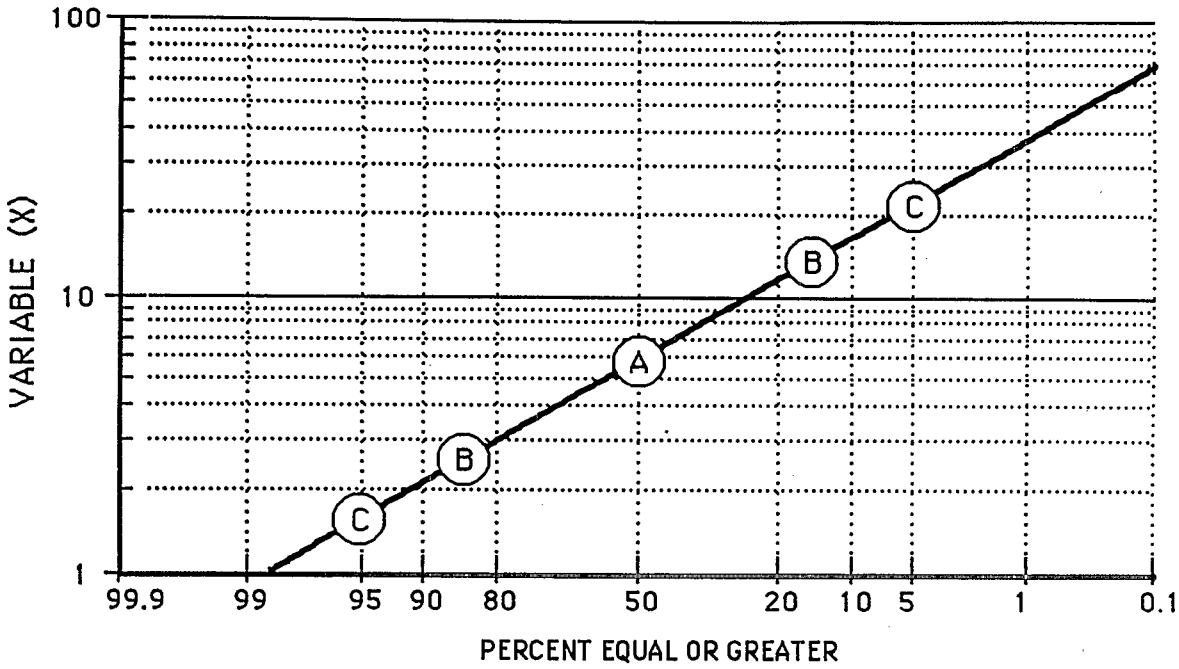
Once the appropriate statistical parameters have been computed, the expected frequency at which a specific stream concentration of interest will occur can be determined in one of two ways.

A probability distribution plot can be constructed on log probability paper, using the estimated statistical parameters for the in-stream concentrations. Using this procedure, the probability of occurrence of the concentration of interest can be scaled directly from the plot. An example of this approach is illustrated by Figure 4-2.

Alternatively, the concentration that will not be exceeded at some specific frequency (or probability) can be calculated using Equation 3.9. Similarly, this equation can be used to estimate the stream concentration for a given probability of exceedance by using ($1-\alpha$) for the calculation.

The probability relationship shown on Figure 4-2, or specific-percentiles calculated by Equation 3.9, represents the distribution of stream concentrations during precipitation related wet periods, i.e., when urban runoff takes place. The concentration shown is the result of both upstream contributions and urban runoff discharges. The distribution of stream concentrations will be lognormal when computed by the moments approximation illustrated here. As indicated in the sensitivity analyses presented later, this is an approximation of the actual stream concentration distribution.

The probability of exceeding a specified concentration developed under the previous step may be converted to a mean recurrence interval (MRI) by taking the reciprocal of the probability value. However, because runoff flow and concentration parameters are based only on conditions when a runoff event is in progress, the recurrence interval computed above is in terms of events. For example, a 1.0 percent exceedance probability ($P = 0.010$)



A lognormal distribution plots as a straight line on LOG-PROBABILITY graph paper

Only two of the points shown are required to define the line

(A) The median is plotted at the 50 % probability $\tilde{X} = EXP(\mu_{ln}(X))$

(B) The 16th and 84th percentile ($\pm 1\sigma$) plotting positions are determined as

EXP(log mean $\pm$ log std dev)

(C) The 5th and 95th percentiles are determined using $Z = \pm 1.645$ in the following equation

$$X = EXP(\text{log mean} + Z * \text{log std dev})$$

Other plotting positions can be computed using appropriate Z scores which correspond with the selected percentiles

Figure 4-2. Construction of Probability Distribution Plot

has a mean recurrence interval of 1/0.010 = 100 storm events. On average, the concentration corresponding to this probability will be exceeded once in every 100 storm events.

It is often desirable, however, to estimate the recurrence interval in time units, to aid in evaluating the significance of the conditions produced by urban runoff. This conversion can be made using the average number of storm events per year determined from analysis of rainfall records as described in Chapter 3. As a first order approximation, Equation 3.43 may be used. For example, if the local area has an average of 100 storms per year, then the selected concentration that has a 1.0 percent exceedance probability (1 in 100 storm events) will be exceeded on an average of once per year.

This first order approximation is a simplification of the recurrence pattern that is likely to result in natural systems. Multiple exceedances would be expected in some years, and none in others, with the clusters of exceedance events recurring at some interval greater than the MRI of 1 year produced by the first order approximation. If this is kept in mind when interpreting and comparing results, the MRI can be used to provide a relevant basis for evaluating the results.

4.4 ILLUSTRATIVE EXAMPLE OF ANALYSIS

A step-by-step application of the Analytically Derived Distribution Approach method is presented in Appendix C to further assist the user in understanding this methodology. This illustrative example also indicates how results of this analysis may be interpreted to guide evaluations of the significance of urban runoff water quality impacts and decisions on control actions.

Appendix C includes a set of input data used for the example analysis, the computations and interim results. The sample examines the effect of urban runoff pollutant sources by setting upstream concentration at zero. A comparison between these results and results obtained using the Direct Numerical Approach and the Monte Carlo Simulation technique for a variety of input conditions is presented in Section 4.5.

The selected input parameters for the example have been developed based on the following assumptions concerning the original sources of data:

- Urban Runoff Quality: Copper is selected for the analysis. Owing to limited local monitoring data the best estimate of mean and variability of copper concentrations is based on the NURP estimate (WCC 1983) of the quality characteristics of the median urban site.

- Streamflow: It is assumed that there is no USGS flow gage available immediately upstream of the urban area to estimate the streamflow characteristics. However, analysis of flow records for appropriate gaged streams in the area support an estimate of the average streamflow of 1.25 cfs/square mile of drainage area, with a coefficient of variation based on daily flow values of 1.25. Total drainage area upstream of the urban area contributing runoff to the stream is 100 sq.mi. Thus, the average upstream flow is (100 sq. mi. * 1.25 cfs/sq. mi.) 125 cfs. This value is used as an input for the computations.

- Runoff Flow: Assume that the urban area under consideration has an area of 1 sq.mi., a runoff coefficient (Rv) of 0.25, and that long-term rainfall records for a local gage have the following characteristics (based on statistical analysis using the SYNOP program WCC (1988).

Parameter	Units	Mean	Coefficient of Variation
Storm average intensity	[in/hr]	0.075	1.35
Storm volumes	[inch]	0.40	1.50
Storm durations	[hour]	7.0	1.20
Storm intervals	[hour]	87.6	1.00

Using the above data, the average runoff rate for the 1 square mile of urban area is computed using the mean storm intensity and the runoff coefficient

$$(0.075)\ (0.25)\ (645.33) = 12 \text{ cfs / sq. mi.}$$

where the factor 645.33 converts inches per hour to cfs / square mile.

The coefficient of variation of runoff flow (Q_R) is assigned a value of 1.15 which is 85% of the coefficient of variation for the storm average intensity.

These values are listed as the input parameters on Appendix C. The subsequent computations to determine the probability of resulting in-stream concentrations during storm events are shown on Appendix C. Figure 4-3 is a plot of the output. The estimated probabilities of in-stream concentrations are then converted to average recurrence intervals to provide a more understandable basis for interpreting and evaluating results. Figure 4-4 presents the average recurrence interval associated with specific stream concentrations that will be produced during storm runoff periods.

4.5 COMPARISON OF THE THREE PROBABLISTIC METHODOLOGIES

All three models were run using the data presented in Section 4.4. The results from each model are presented in the form of the CDF of downstream concentrations. In addition,simulations were made using a wide range of possible values for the inputs. Results of these simulations are presented in tabular form.

Figure 4-5a shows the CDF for downstream concentration resulting from simulations conducted using all three probabilistic models. At higher cumulative probabilities (greater than 90%) all three methods result in approximately the same downstream concentrations. This is clearly shown in Figures 4-5 a and b. The latter presents only the highest 10% of the downstream concentrations. At lower probabilities the Analytically Derived Distribution Method deviates from the other two solutions. This is due to the choice of the 90 and 99.6 percent quantiles for estimating the parameters of the assumed log normal distribution for the dilution. These percentiles were selected since the intent here is to estimate downstream concentrations with return periods of the order of 3 years or so that generally correspond to concentrations above the 90th percentile (see Section 3.3). Note that by varying these percentiles, the Analytically Derived Distribution Approach can be forced to agree with other sections of the CDF.

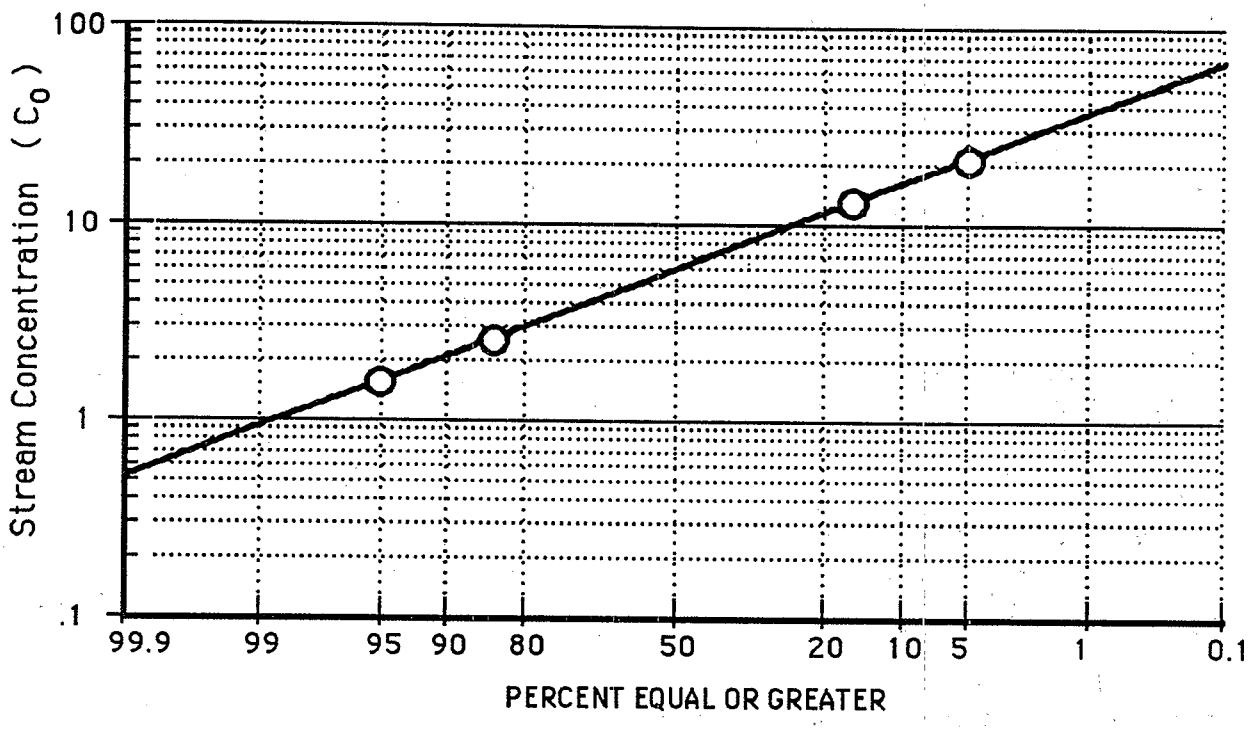


Figure 4-3. Probability Distribution of Pollutant Concentrations During Storm Runoff Events

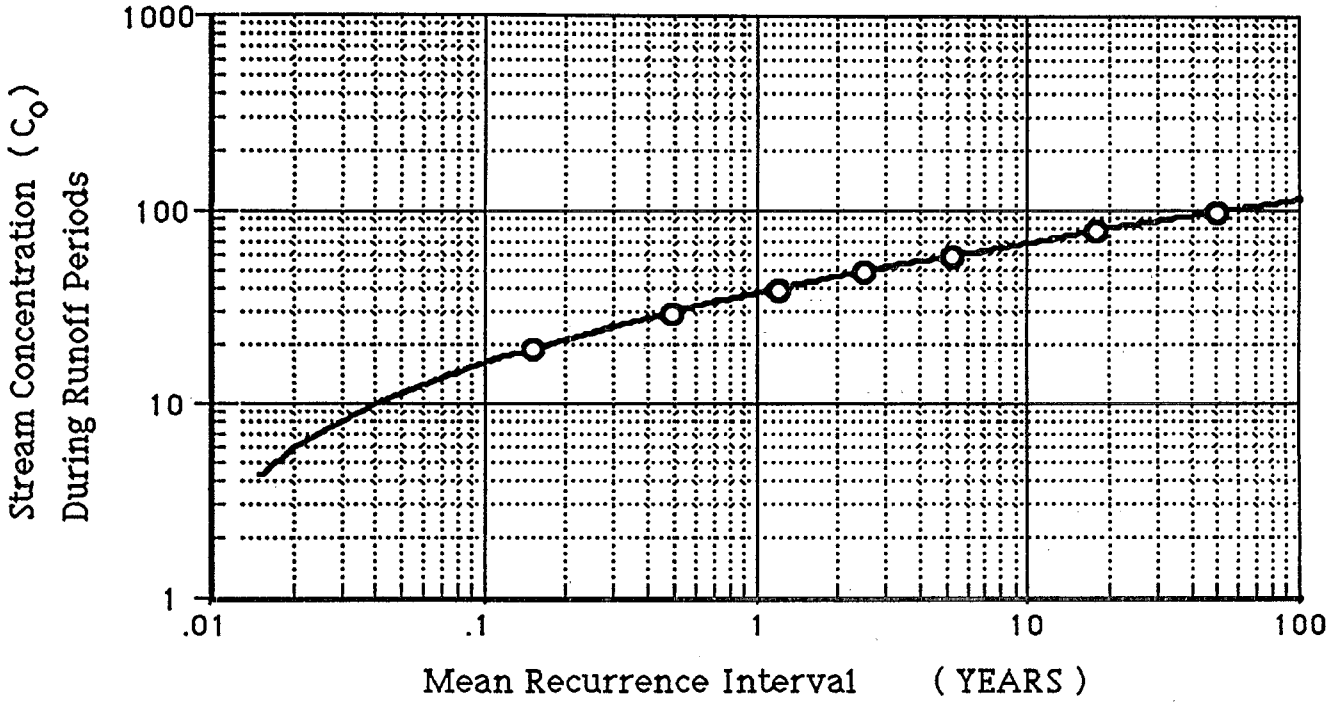


Figure 4-4. Recurrence Intervals for Stream Concentrations

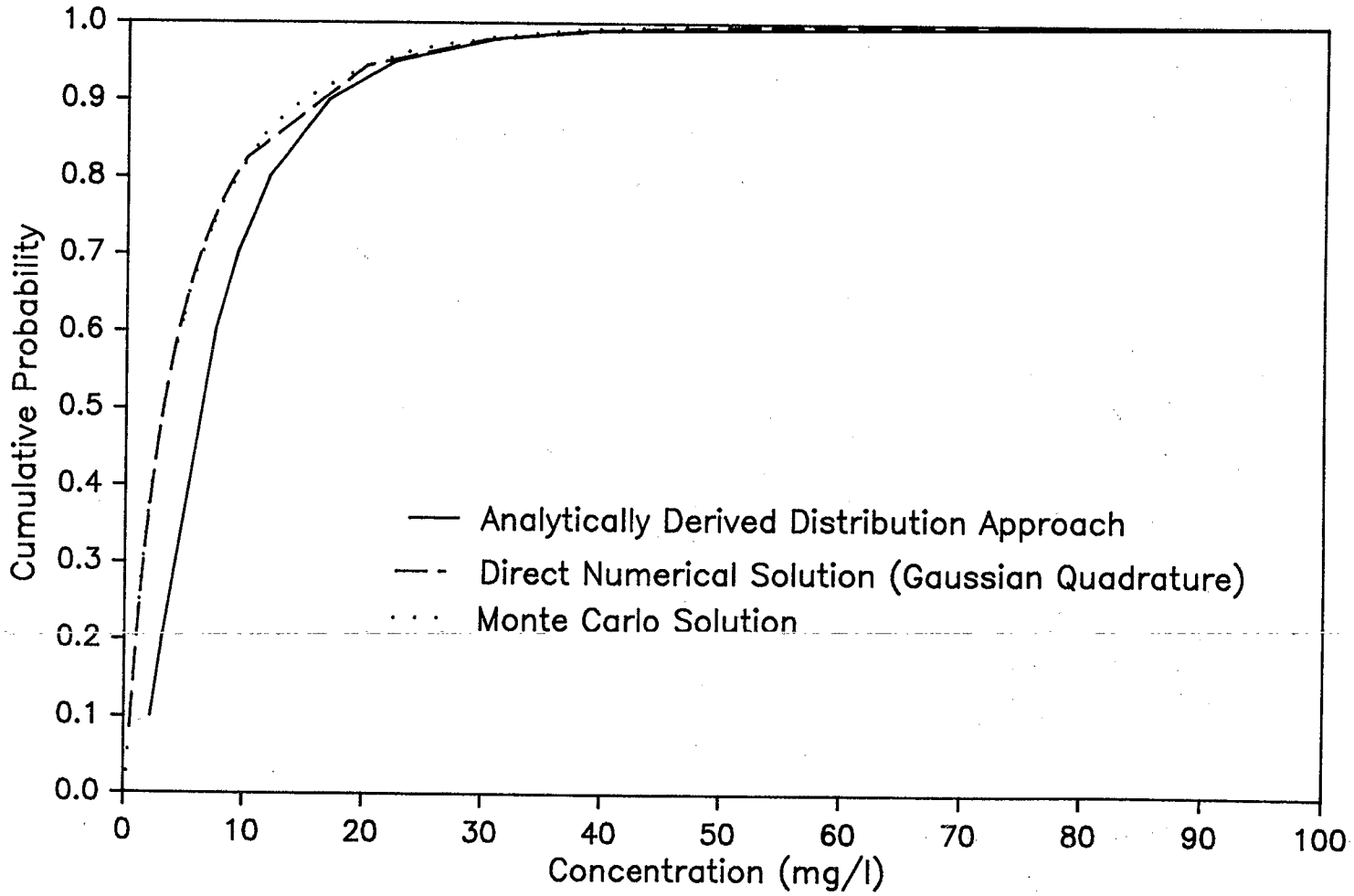


Figure 4–5a. Comparison of CDFs for Three Probabilistic Methodologies

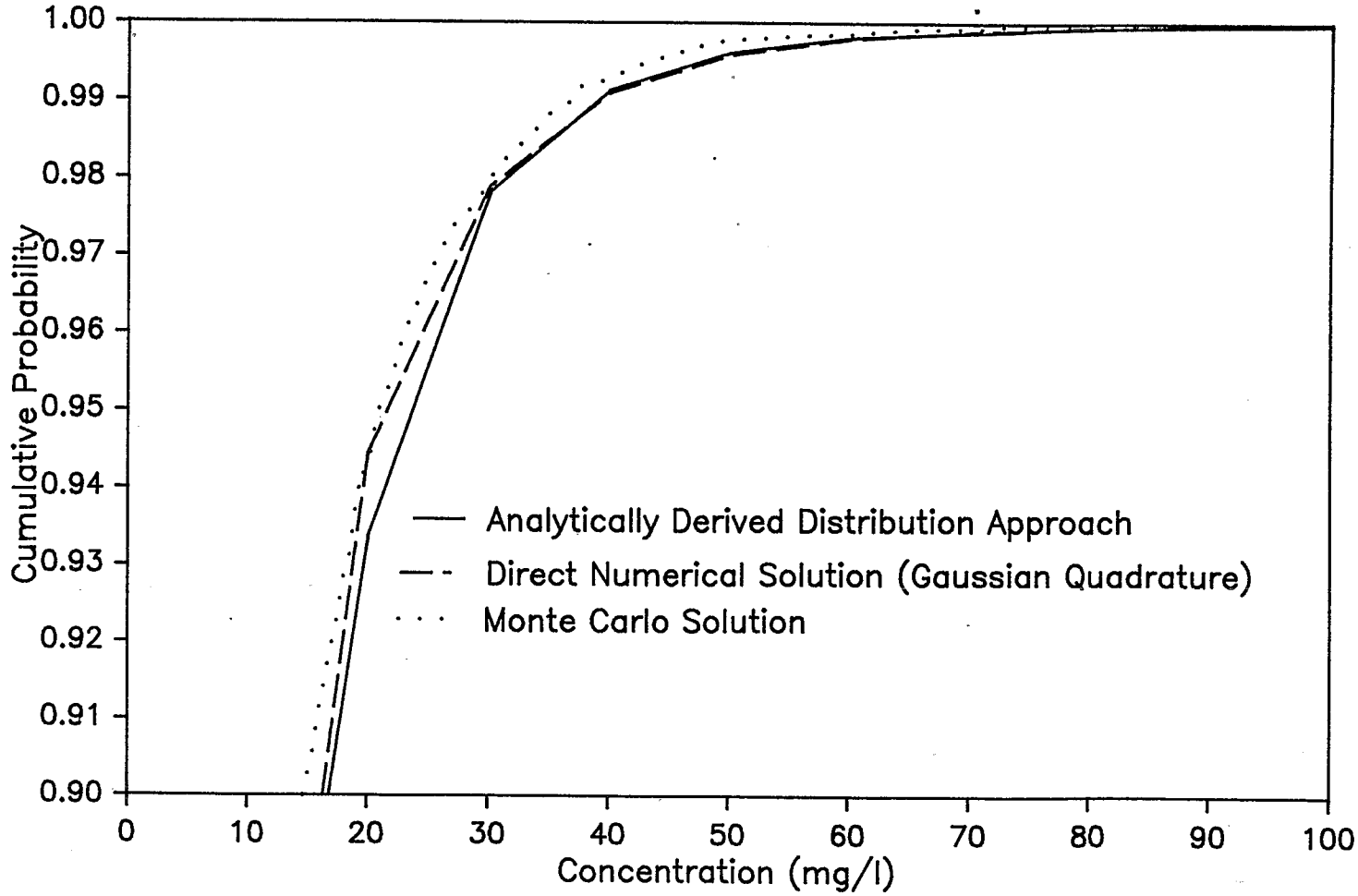


Figure 4–5b. Comparison of CDFs for Three Probabilistic Methodologies (highest 10% of values)

A series of simulations were conducted using (i) two different upstream concentrations, (ii) four different ratios of runoff flow to streamflow, and (iii) two different ratios of coefficients of variation of upstreamflow to runoff flow so that the three methods could be compared over a wide range of possible inputs. Table 4-2 presents the parameters for the model input distributions. Since the solution of the model is primarily dependent upon the ratio of runoff flow to upstream flow, only the upstream flow was varied. Results from the simulations are presented in Table 4-3.

A comparison of Run 1 to Run 2 shows the effect of increasing the upstream concentration from essentially zero to one half of the runoff concentration. The maximum difference in the downstream concentrations between the three methods was 7.8 mg/l at the 99 percentile value. This is less than about 15% of the mean concentration for the three methods and is within the accuracy range for a screening level methodology. For the other percentiles, the maximum difference between methods is even less.

Comparison of Runs 3 and 4 show the effect of changing the ratio of upstream flow to runoff flow from 10 to 100 using a ratio of coefficient of variation of 2. Runs 5 and 6 use the same flow ratios but have a ratio of coefficients of variation of 3. In general, the three methods produce essentially identical results using the above inputs. However, for the smaller flow ratio with a higher ratio of coefficient of variation (Run 5) the Analytically Derived Distribution starts to deviate from the other two methods. For percentiles from 90 to 99.5 the Analytically Derived Distribution Approach over estimates the concentration by 14% to 30% relative to the Monte Carlo solution. The larger differences occur at the lower percentiles. The Direct Numerical Solution and the Monte Carlo Method are usually within 10% of each other.

Runs 5 through 8 show the effect of changing flow ratio only. For these runs the flow ratio varies from 1 to 1000. The results from these runs indicate that as the flow ratio decreases the Analytically Derived Distribution Approach predicts higher concentrations relative to the other two methods. If the runoff flow equals the upstream flow, the Analytically Derived Approach can overestimate the concentration by over 40% relative to the Monte Carlo Method and 25% relative to the Direct Numerical Solution Approach. The Direct Numerical Solution Approach deviates from the Monte Carlo Method solution at these flows also. For Run 7 it predicts concentrations about 15% greater than the Monte Carlo Method.

Thus, for most sets of input parameters, the three methods estimate approximately the same answer. However, when the upstream flow and runoff flow are of about the same magnitude, the Analytically Derived Distribution Approach generally yields higher estimates of concentration relative to the other two methods.

4.6 INTERPRETATION AND EVALUATION OF RESULTS

The significance of a particular magnitude/frequency of stream concentration caused by urban runoff can be evaluated by comparing it with concentrations that are relevant for the beneficial uses of the water body. If, for example, protection of aquatic life is the predominant beneficial use of the urban stream, the concentrations that result from urban runoff may be compared with a stream target concentration associated with appropriate water quality criteria or standards.

Figures 4-6 and 4-7, which repeat the probability and recurrence interval plots developed for the example analysis, illustrate how analysis results may be used for NPS planning evaluation.

Table 4-2
Inputs Used in Sensitivity Analysis of the PDM Model

Run	CS		CR		QS		QR	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	0.01	0.01	40.8	24.48	125	156	12.0	13.8
2	20.4	12.24	40.8	24.48	125	156	12.0	13.8
3	0.01	0.01	40.8	24.48	10.0	15.	1.00	0.75
4	0.01	0.01	40.8	24.48	100.	150.	1.00	0.75
5	0.01	0.01	40.8	24.48	10.0	30.	1.00	0.75
6	0.01	0.01	40.8	24.48	100.	300.	1.00	0.75
7	0.01	0.01	40.8	24.48	1.0	3.	1.00	0.75
8	0.01	0.01	40.8	24.48	1000.	3000.	1.00	0.75

Table 4-3
Comparison of Results between Monte Carlo Method, Analytically Derived Distribution Approach and Direct Numerical Approach

Run	Method	Mean	Median	SD	Percentiles 90%	95%	99%	99.5%
1	MC	5.93	3.2	7.63	14.6	20.8	36.8	43.2
	Moments	8.3	3.31	7.8	16.8	22.4	38.3	46.7
	Quad				14.7	21.0	38.0	48.2
2	MC	23.1		11.6	38.0	44.9	60.8	67.7
	Moments	25.0	22.4	12.0	39.6	46.8	63.7	71.3
	Quad				39.4	47.4	68.6	73.9
3	MC	7.36	4.2	8.68	17.9	24.3	40.7	47.6
	Moments	11	4.2	8.8	20.7	26.9	44.2	52.9
	Quad				17.7	24.7	43.6	53.0
4	MC	1.18	0.50	2.08	2.9	4.7	9.5	12.2
	Moments	1.3	0.62	2.3	2.9	4.5	10.1	13.6
	Quad				2.8	4.5	10.2	13.6
5	MC	11.6	6.69	13.6	29.1	38.3	62.5	72.8
	Moments	21	17.3	14	37.7	47.0	71.1	82.7
	Quad				28.0	38.1	63.8	76.1
6	MC	2.81	0.85	5.31	7.57	12.7	26.1	32.3
	Moments	3.3	1.74	5.3	7.4	11.1	24.0	31.8
	Quad				7.0	11.5	25.7	33.2
7	MC	23.0	18.6	17.6	45.6	56.6	85.2	97.6
	Moments	38	32.1	23	65.5	80.1	117	134
	Quad				52.1	65.1	98.6	114
8	MC	0.434	0.10	1.40	0.96	1.77	5.23	8.56
	Moments	0.39		1.40	0.85	1.53	4.60	6.89
	Quad				0.88	1.53	4.69	6.57

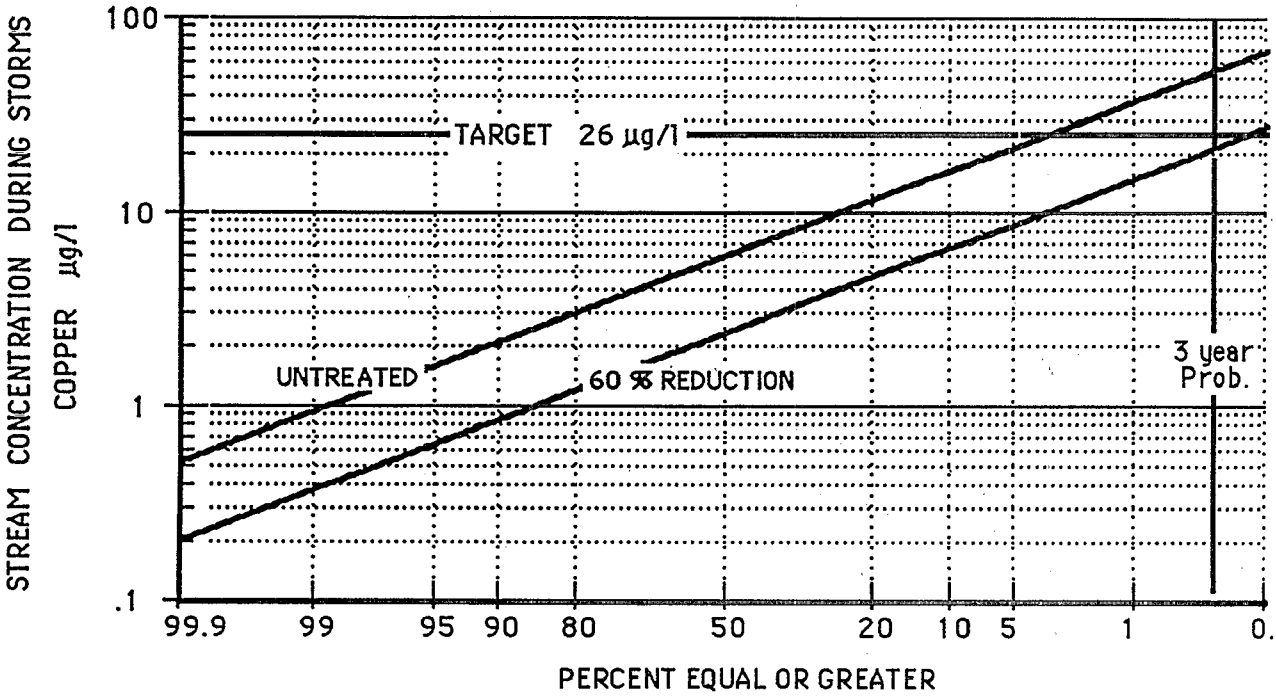


Figure 4-6. Use of the Methodology to Evaluate the Water Quality Impact of an Urban Runoff Control Program (1)

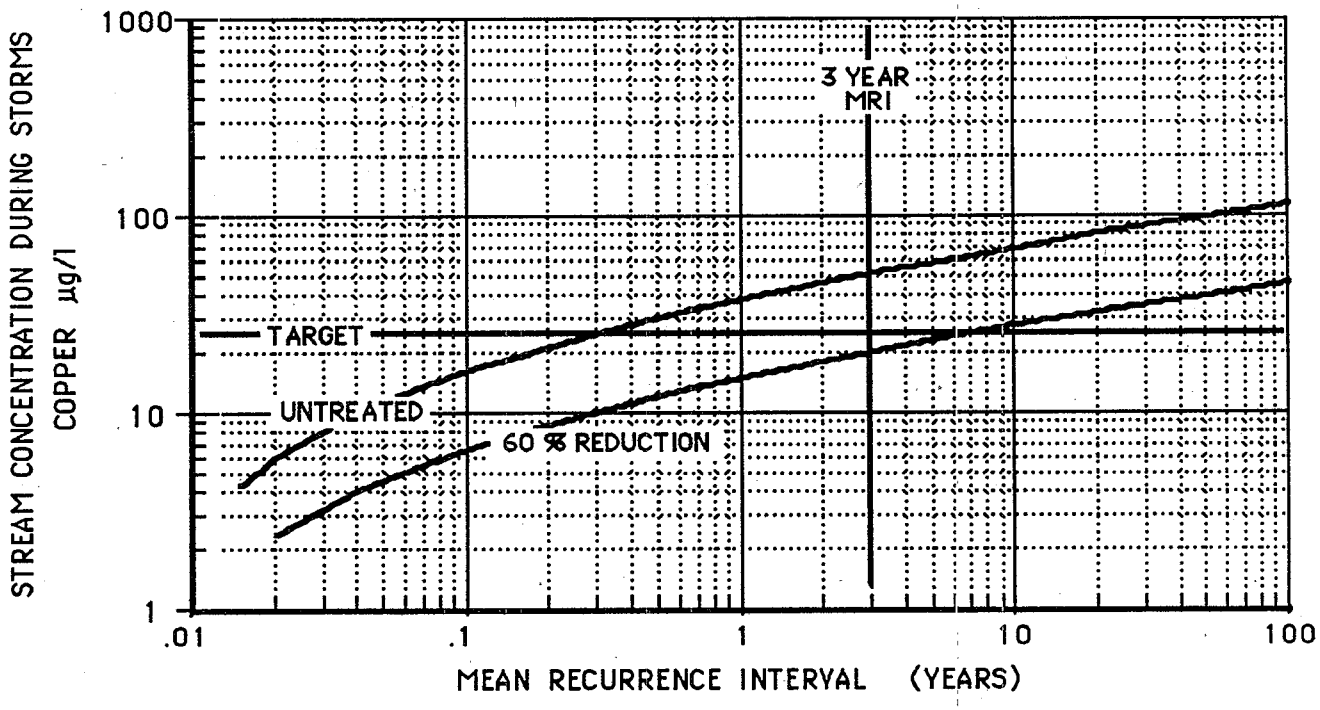


Figure 4-7. Use of the Methodology to Evaluate the Water Quality Impact of an Urban Runoff Control Program (II)

If the runoff concentrations used in the analysis were for copper,the projected stream concentrations can be compared with the EPA toxicity criteria for copper, by showing this value on the concentration axis. The acute toxicity value is selected as most appropriate for comparison with the effect produced by intermittent short duration exposures. If the total hardness of the receiving stream is 150 mg/1, the copper criterion value is 26 µg/l, and a horizontal line is drawn on the plot at this location. This is considered to reflect a safe level when it is exceeded no more frequently than once in 3 years. This aspect of the target is designated on the plot by a vertical line locating this frequency of occurrence, at 3 years on the MRI plot. For the number of storms per year used in the example, this corresponds to an exceedence probability of 0.33 percent on the probability plot.

In this case, the stream concentrations of copper caused by untreated urban runoff discharges exceed the "target" to a significant extent. The target concentration is exceeded at a recurrence interval of about 0.3 years (every 100 days on average), and at a 3-year recurrence, the stream concentrations produced during runoff events will be about double the target.

The water quality effects of proposed alternative NPS control programs can be assessed by repeating the analysis using revised input values to represent the characteristics of the urban stormwater runoff. To illustrate, it is assumed that an NPS control program that might be implemented could effect a 60 percent reduction in the mean copper concentration in the urban runoff. Results of the analysis are plotted on the same figure to provide a graphic display of the degree of mitigation of water quality impacts that such a program is projected to produce. In this example, the 3-year concentration is reduced to about 20 µg/l, safely below the target. The target concentration of 26 µg/l is now projected to be exceeded not more than once every 7 years on average.

Such an approach is useful in the evaluation of the significance of urban runoff and for identifying potential water quality effects from the implementation of control measures. There are two reasons why the use of a probabilistic approach is important. First, because of the stochastic nature of the urban runoff flow system, virtually any target concentration will be exceeded at some frequency, however small. Second, from a practical point of view, there are limits to the capabilities of controls. If the receiving water total hardness in the example was 50 mg/l, the criteria concentration would be 9 µg/l rather than 26 µg/l, and the improvement produced by the control would fall far short of meeting the targets. Receiving water impacts may be significantly reduced, but will never be totally eliminated. Therefore it is necessary to evaluate the degree of water quality improvement associated with different control alternatives and to implement control measures to attain water quality standards.

CHAPTER 5.0

EVALUATION OF METHODOLOGY

5.1 INTRODUCTION

This chapter evaluates several different aspects of the application of the probabilistic methodology. Section 5.2 presents a verification analysis that compares computed and observed in-stream concentrations of pollutants during storm runoff periods. Section 5.3 presents indirect evidence of the method's reliability by comparing the projected use impacts, inferred from computed concentration probability distributions, with local observations relating to protection or impairment of the use at particular sites. Finally Section 5.4 discusses the results of sensitivity analyses performed to evaluate the significance of uncertainties or alternate estimates for values of specific input parameters.

5.2 VERIFICATION OF COMPUTED CONCENTRATIONS

Using site-specific data on stream flows and concentrations, and urban runoff flows and concentrations, acquired during the NURP program (WCC 1983), a test of the reliability of the methodology has been performed. Data from the Rapid City, SD project, shown schematically on Figure 5-1, are used.

Data from approximately 20 storm events were available, which provided event mean values for stream flow, runoff flow, and concentrations of four pollutants in the urban runoff and in the stream above and below the discharge point during the storm events. The four pollutants included Total Suspended Solids (TSS), carbon oxygen demand (COD), total phosphorous (P), and lead. Since the selected monitoring site for urban runoff enters Rapid Creek downstream of stream monitoring stations, the urban runoff flow and quality characteristics, derived from data collected at Station #6, were assumed to apply for an urban catchment closely similar in area and land use distribution, which discharged between stream monitoring stations at which concurrent data were available.

For each of the data sets, consisting of about 20 event mean values for stream flow (Q_S), upstream concentration (C_S), runoff flow (Q_R), and runoff concentration (C_R), the statistical properties of the log transforms were estimated. The calculated probability distribution of instream concentration was then compared with the distribution of measured concentration at downstream stations. This information is presented as discussed below:

- Table 5-1 summarizes the statistical characteristics of the assumed log normal distributions of the monitored events, and presents the results computed using the Analytical Derived Distribution Method for the four pollutants.
- Figure 5-2 shows the probability distribution of streamflows and runoff flows for the 20 monitored events, and the log-normal distribution assigned to the observed data. Two upstream stations are considered individually and the data from each are shown.
- Figure 5-3 presents a comparison of concentrations for TSS in runoff and upstream locations during monitored events. It also compares the computed probability distribution of the in-stream concentration of TSS and the observed data. An

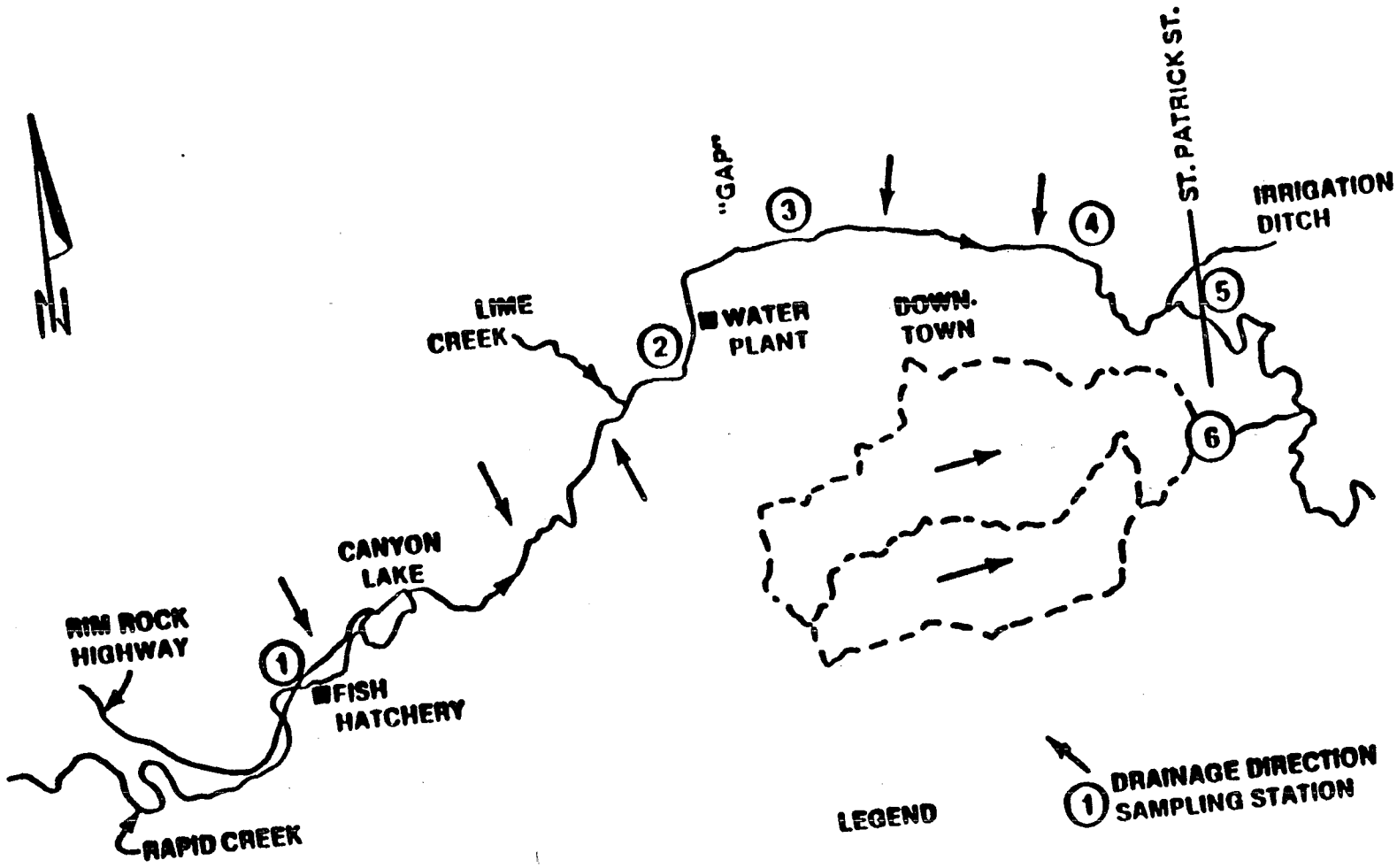


Figure 5-1. Map of Rapid Creek and Major Drainage Basins within Rapid City Proper; Sampling Site Locations

TABLE 5-1. RAPID CITY VERIFICATION TEST

Item		Mean	Median	Coefficient of Variation
Stream Flow (CFS)	Q_S	104	90.5	0.57
Runoff Flow (CFS)	Q_R	17.5	12.6	0.96
Runoff Conc. C_R	TSS	3689	2760	0.89
	COD	219	195	0.52
	TP	2198	1435	1.16
	Pb	382	253	1.13
Upstream Conc. C_S	TSS	325	158	1.80
	COD	30	25	0.66
	TP	206	100	1.80
	Pb	19	4	4.70
Calculated Downstream Conc.	TSS	822	523	1.21
	COD	58	47	0.74
	TP	500	293	1.38
	Pb	73	35	1.79

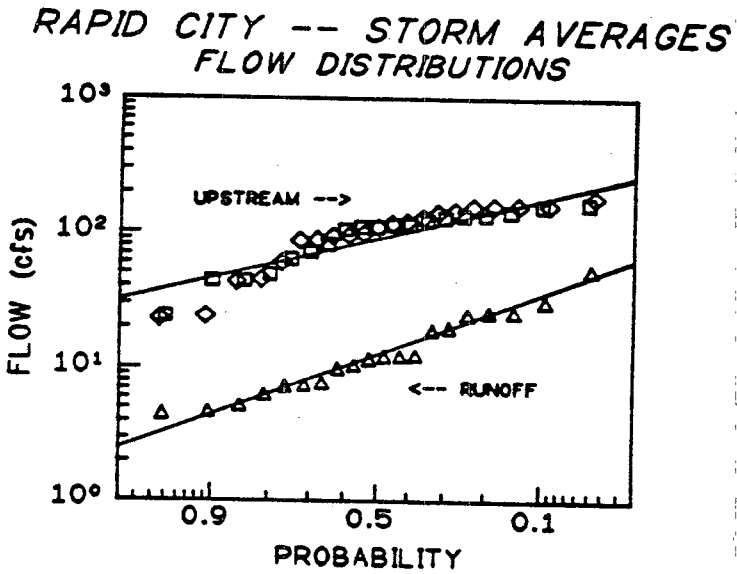
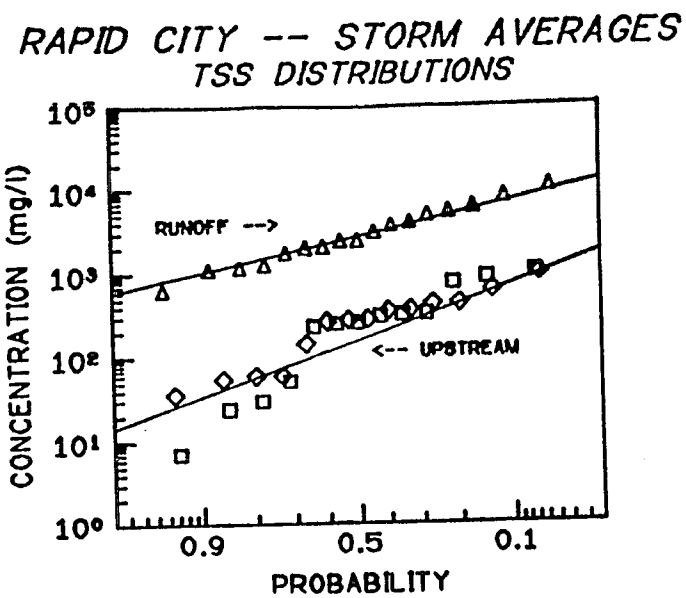


Figure 5-2. Probability Distribution of Stream and Runoff Flows



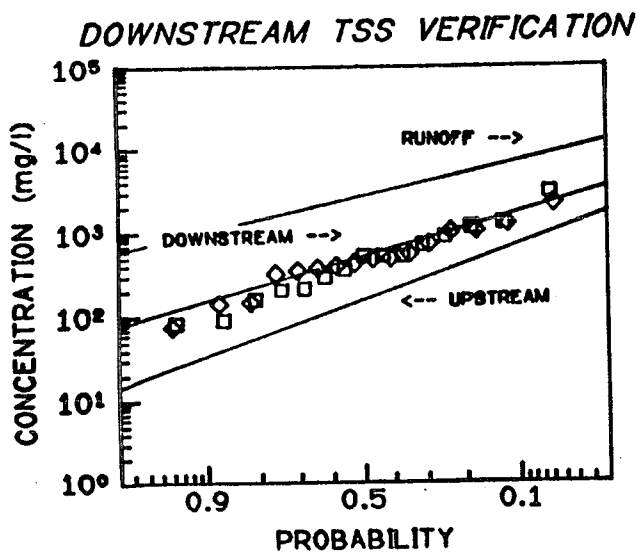
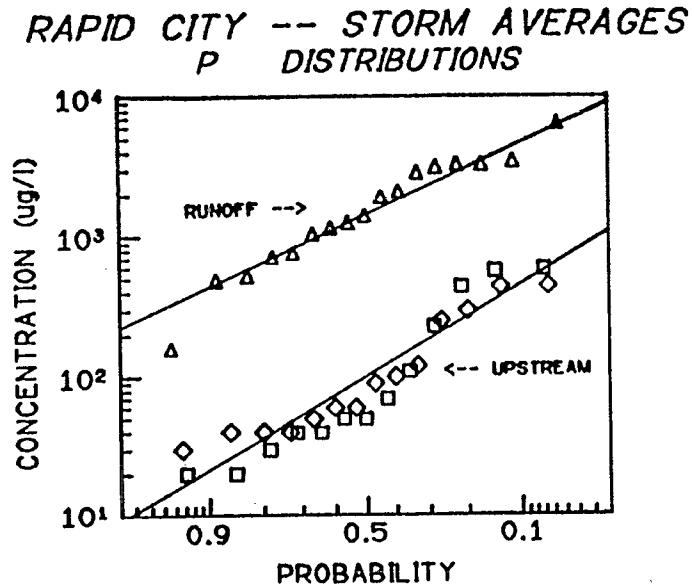


Figure 5-3. Comparable Information on Concentrations in Runoff and Upstream Locations and the Probability Distribution



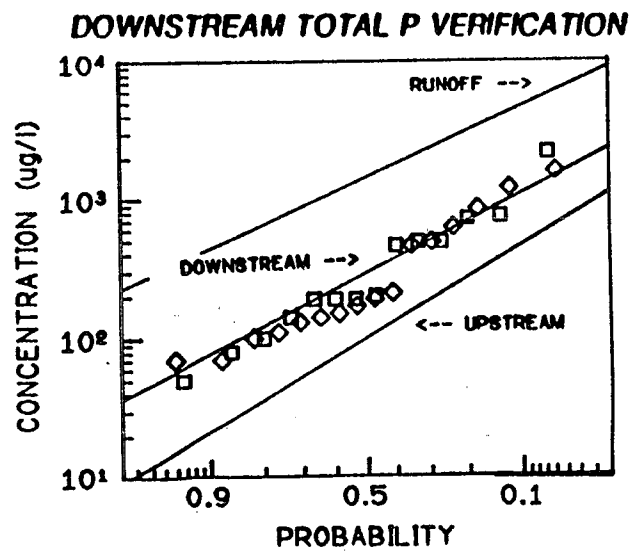


Figure 5-4. Results for Total P

excellent match between the data and the calculated distribution supports the methodology used.

- Figures 5-4, 5-5, and 5-6 present results for Total P, COD, and Lead, respectively. In each case, the model results compare well with the data.

The results and comparisons presented above are based on an available "sample" of storm events and are representative of the conditions which prevailed during the monitoring activities. As such, this analysis is appropriately a "test" of the analysis procedure, and does not necessarily reflect the projected long term conditions.

The above verification tests confirm the validity of the computational methods. These results are essentially a test of the methodology because the data used represent a limited sample of overall conditions drawn from monitoring results. Such conditions may, and often do, represent a biased representation of typical long term conditions. The favorable comparisons, however, do provide confidence that the long term projections used to guide decisions will be sensitive to how well the long term input characteristics can be defined.

5.3 VERIFICATION OF IMPLIED EFFECT ON BENEFICIAL USE

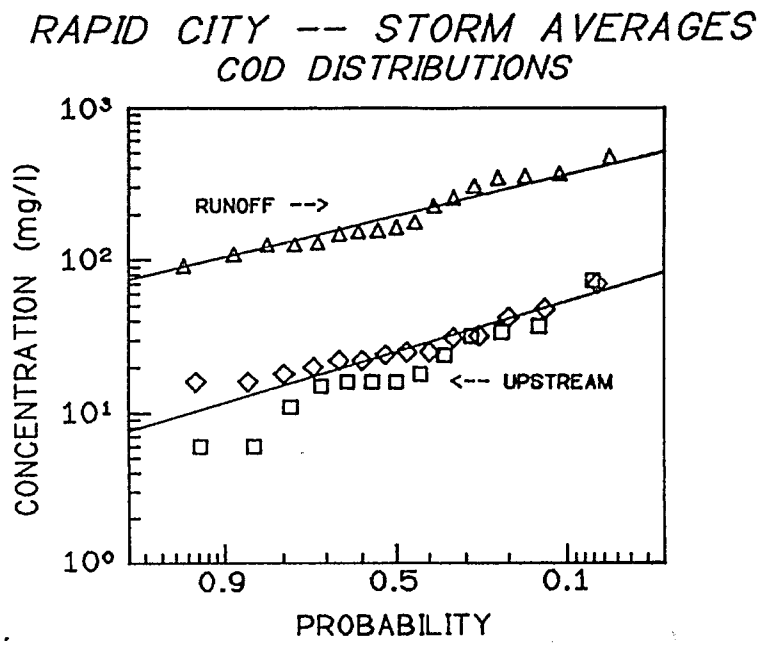
A second basis for determining the reliability of the methodology, in identifying the potential of urban runoff discharges to cause significant impairment of a receiving water body, is to compare the expected impact levels with observed biological conditions. This requires (i) sufficiently accurate estimates of input parameters, (ii) accurate estimate of in-stream concentrations, and (iii) the resultant exposure concentrations and the biological effects are accurately evaluated.

The information available to examine this issue is more tenuous than that available for checking the computation procedures of the methodology. Three of the NURP projects examined the effects on aquatic life in streams receiving runoff from monitored sites:

- Bellevue, WA, concluded that whatever adverse effects were observed were attributable to habitat impacts (streambed scour and deposition) as opposed to chemical toxicity. For this project, heavy metal concentrations in the monitored urban runoff sites were typical of the average for all urban sites. The screening analysis projections under these conditions did not indicate the potential for a water quality problem.

- Tampa, FL, conducted extensive bioassay tests but failed to show any adverse effect of pollutants in urban runoff. When the screening analysis was implemented using site concentrations representative of Tampa monitoring results, a toxicity problem was not predicted.

- Rapid City, SD, conducted extensive biological testing of phytoplankton, periphyton, benthic organisms, and fish in the stream receiving urban runoff from the entire city. Observable adverse effects were not shown to exist nor were any significant incremental changes demonstrated as the river picked up an increasing component of urban runoff. A healthy cold water fishery was reported to exist in the stream. The screening analysis indicated that a significant impairment would not be anticipated, even though streamflows tend to be relatively low and monitored lead concentrations in urban runoff are at the higher end of the range for urban sites.



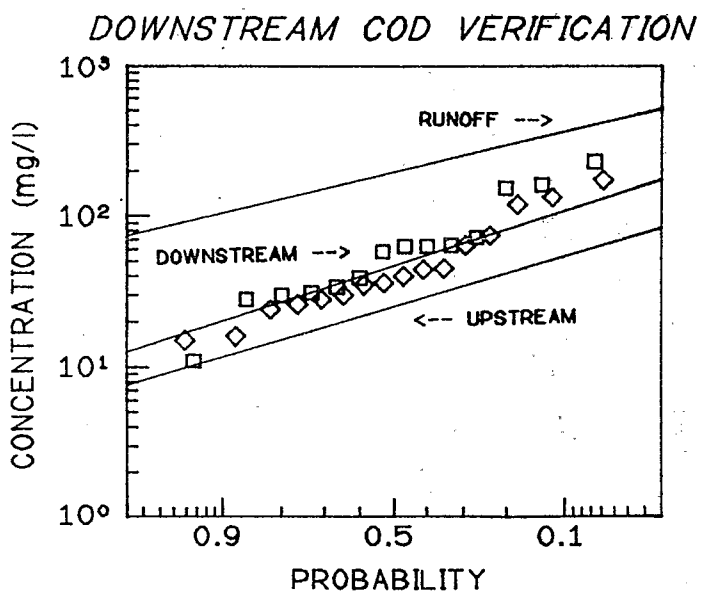
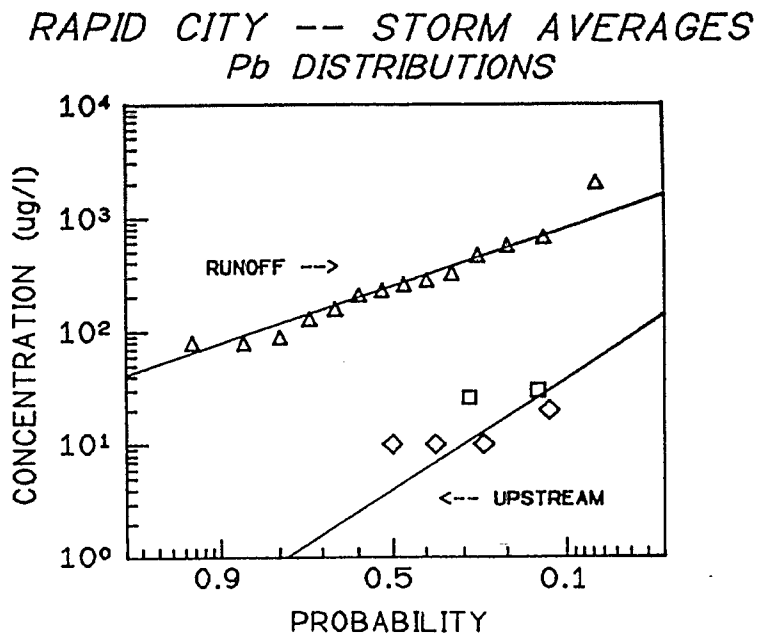


Figure 5-5. Results for COD



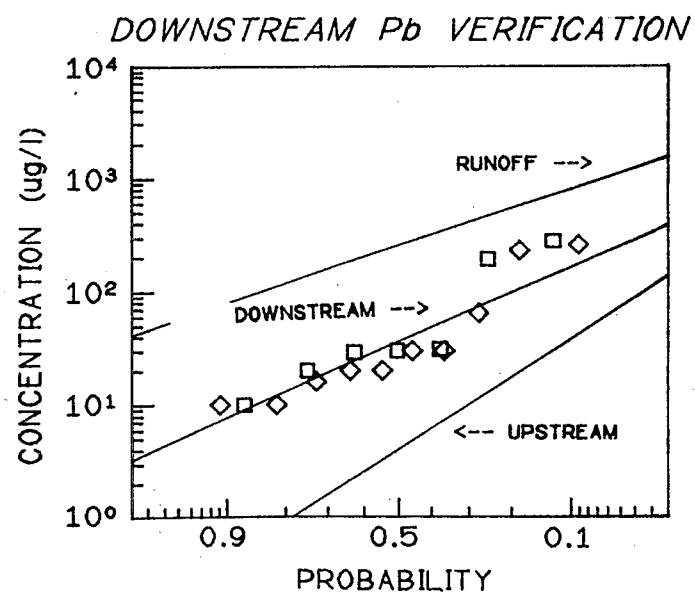


Figure 5-6. Results for Lead

5.4 SENSITIVITY ANALYSES

This section presents a series of comparisons between the results obtained under different sets of conditions. The intent is to provide information on the sensitivity of model results to certain choices and assumptions that a user might have to consider due to uncertainty concerning the input parameters. The following issues are addressed:

- Analytically Derived Distribution Approach versus the Direct Numerical Approach(quadrature)
- Effect of correlation between stream and runoff flows
- Effect of nonzero upstream concentrations

In Section 4.5 the three probablistic methods were compared to each other. A conclusion from this comparison was that the results obtained using the Direct Numerical Approach approximate the results of the Monte Carlo Method. The following sensitivity analysis section presents results obtained using the Analytically Derived Distribution Approach and the Direct Numerical Approach.

5.4.1 Analytically Derived Distribution Approach vs. Direct Numerical Approach

This section examines the accuracy of the approximate solution by presenting a comparison between the Derived Distribution Approach and the Direct Numerical Approach. Results are compared for a variety of inputs. Conclusions are presented on the applicability of the approximate model.

The Derived Distribution Approach offers a number of advantages over the Direct Numerical Approach. The method is simple and hence enables the user to better understand the methodology. A more practical attribute of the method is the speed and ease of use for preliminary screening analyses. It can, for example, be set up in spreadsheet format. This is particularly advantageous when a large number of different conditions are to be examined, e.g., for sensitivity tests or for screening different alternatives. Where the numerical solution computes the probability and return period for a specified stream concentration, the moments approximation additionally allows the user to work in the reverse direction. The ability to estimate the stream concentration that will occur at a user-specified probability or return period proves to be quite useful in many applications.

It has been emphasized that the Derived Distribution Method provides an approximation, and that its projections tend to be conservative. Even allowing for the fact that normal good practice would dictate that final results be confirmed by a more exact computation, it will be useful for the user to develop a sense of the accuracy of the approximate solutions developed by this method. This section compares results obtained by the use of the two computation methods.

An example computation using the Derived Distribution approximation was presented in Chapter 4, and results were summarized on Figures 4-3 and 4-4. The two computation methods are compared for exactly the same input conditions by performing the analysis using the computer program described in Chapter 6. Results are presented and compared with the Derived Distribution approximation in Figure 5-7 which shows the probability distributions, and on Figure 5-8 which shows the mean recurrence intervals.

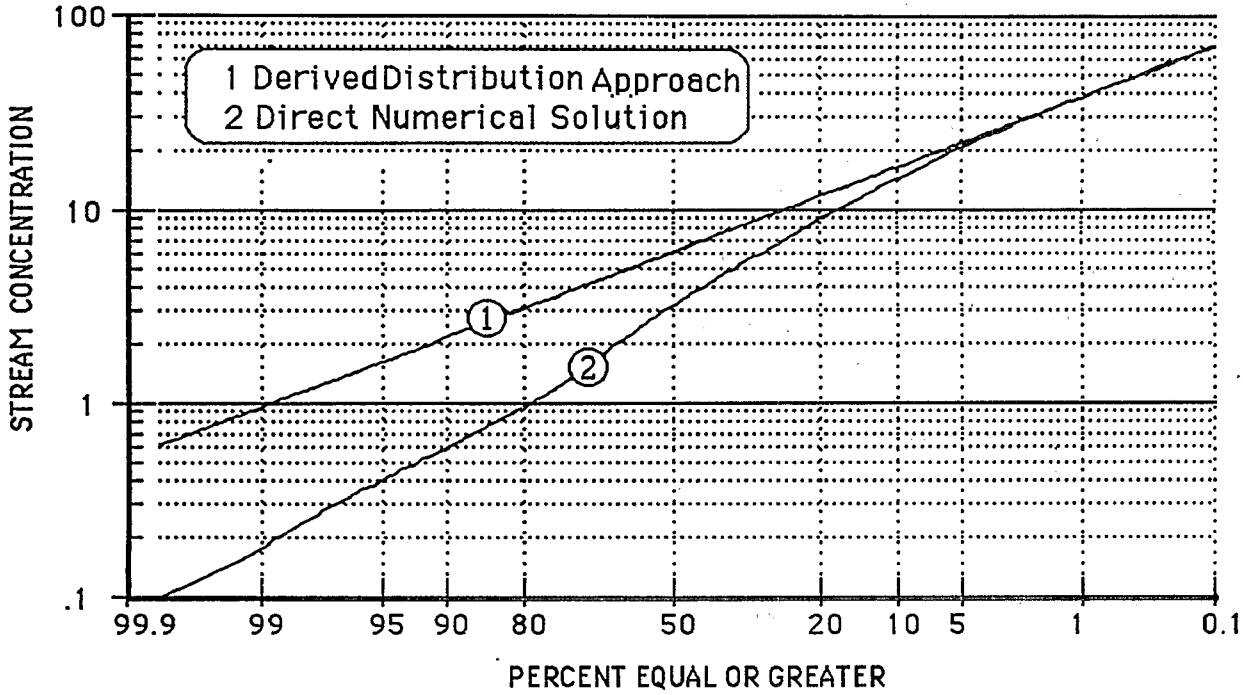


Figure 5-7. Comparison of Results from the Derived Distribution and the Direct Numerical Solution Approachs. Probability Distributions for Selected Input Conditions.

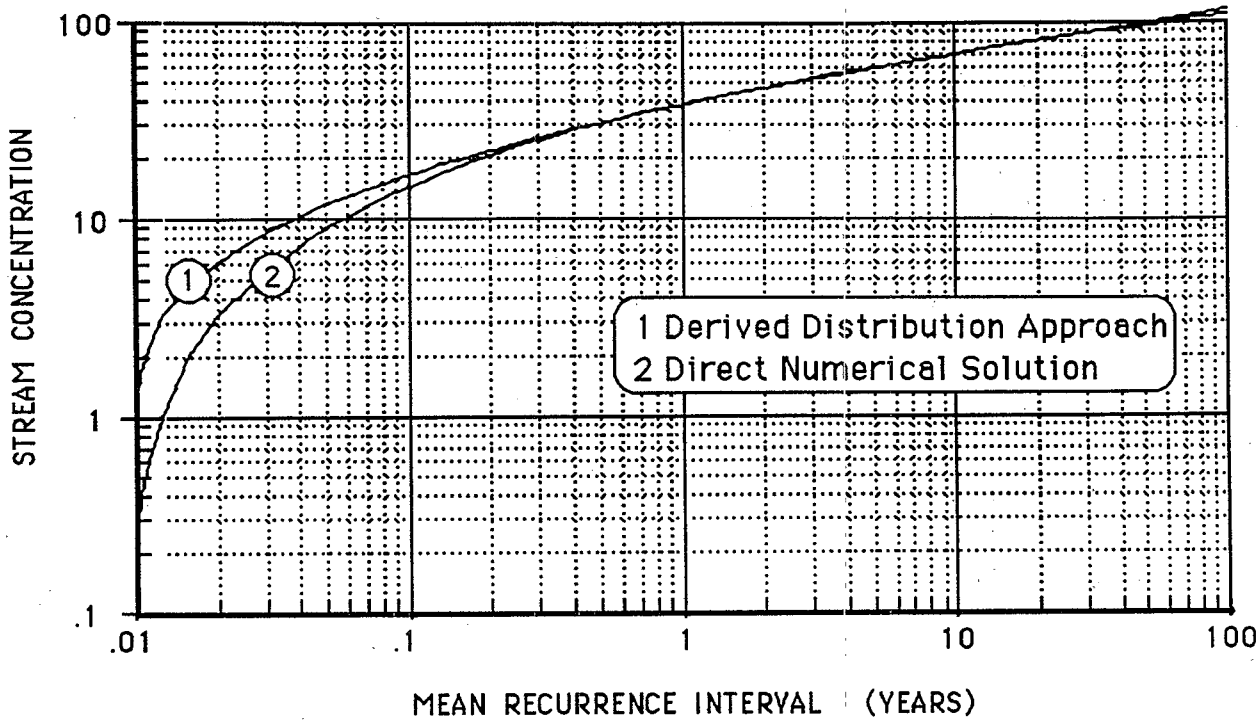


Figure 5-8. Comparison of Results from the Derived Distribution and the Direct Numerical Solution Approaches. Recurrence Intervals for Selected Input Conditions.

Note on Figure 5-7 that the distribution of stream concentrations computed using the numerical method is not log normal. The straight line log normal approximation produced by the Derived Distribution method matches the actual distribution quite well at the upper tail (higher percentiles), but deviates appreciably at lower percentiles. This particular pattern is the result of the selection of the 90th and 99.6th percentiles of the dilution factor (ø) in developing its lognormal approximation. Other input data combinations would modify the match slightly as would the selection of other percentiles for the lognormal approximation of (ø). Sensitivity tests discussed below indicate that the selected percentiles will provide satisfactory matches of the distributions over a wide range of input conditions.

Although the departure of the lognormal approximation over the lower range of probabilities may appear to be a cause for concern, as a practical matter, for the usual water quality impact evaluation it has no effect whatever on any decisions that would be guided by the analysis. Inspection of the probability plot indicates that the moments approximation overestimates the median (50%) stream concentration by 100 percent. But this only represents a concentration of 6 µg/l rather than 3 µg/l. The absolute magnitude of the difference (3 micrograms per liter) is quite small and within the overall accuracy of the method. Further the 50th percentile event has a recurrence interval of about a week in a situation where decisions will be based on recurrence intervals of several years.

The significance of this deviation in terms of its effect on the evaluation of impacts is better shown on Figure 5-8, which compares the recurrence interval projections produced by the two procedures. It emphasizes the close match between the Derived Distribution approximation and the numerical method at MRIs ranging from several months to decades.

The match can be expected to vary depending on the unique combination of input conditions. The relative magnitude of the flows has an important influence, as does both the magnitude and ratio of the coefficients of variation of the inputs. For an assessment of the relationship between outputs of the two computation methods, repetitive analyses were performed, using combinations of input parameters that cover the range of input values that can be reasonably expected to cover most practical situations. The influence of flow rates can be represented by the ratio of mean streamflow to runoff flow (QS / QR), and the range of interest from 1 to 1000. Combinations of coefficients of variation were examined over the range of flow ratios, using values ranging from 0.5 to 3 combined in several different ratios.

Computations of resulting stream concentrations were made using both methods, and the stream concentrations projected to result at a 3-year mean recurrence interval were selected for comparison. The ratio between this concentration computed by the Derived Distribution approximation to that computed by the computer program using the Direct Numerical solution was determined. Results are summarized by Figure 5-9.

The dashed curve shows the comparison that results for conditions likely to be encountered fairly often, coefficients of variation of about 2 for streamflow, of about 1 for runoff flow, and about 0.75 for runoff concentrations. The solid curves define the approximate envelope covering the range of conditions tested.

This sensitivity analysis indicates that the Derived Distribution approximation keyed to the high percentiles of (ø) provides a rather accurate estimate of 3-year stream concentrations, certainly suitable for screening or planning level analyses. Over much of the range of flow conditions, the approximation will provide results within 10% for most of the commonly encountered input combinations. Results will virtually always be expected

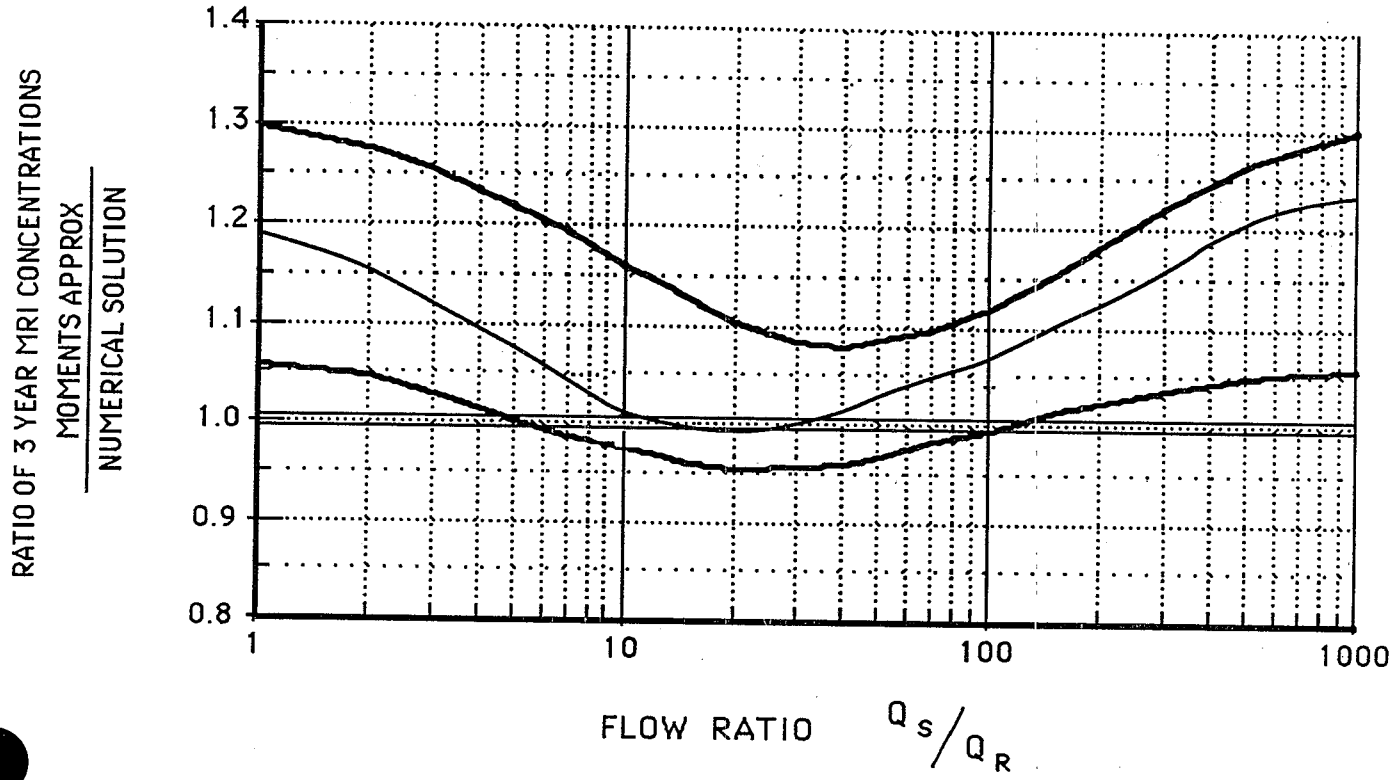


Figure 5-9. Difference in Results between Derived Distribution Approach and Numerical Solution over a Wide Range of Inputs.

to be within 30% , the upper bound being encountered when coefficients of variation and their ratios are high. The Derived Distribution method's estimates of 3-year stream concentrations is seen to be generally conservative, though under some conditions it may underestimate by a few percent.

5.4.2 Correlation between Streamflow and Runoff Flow

This section investigates the effect of a strong positive correlation between streamflow and runoff flow. A positive correlation implies that higher storm runoff events tend to occur when streamflows are at their higher values. This would tend to reduce the stream concentrations that would result from urban runoff. It is important to note that data on correlation coefficients are generally not available.

The Direct Numerical Method was used for computations covering the same range of input conditions examined in the preceding sensitivity analysis. For each combination of input conditions, the projected 10-year MRI concentration was determined for values of correlation coefficient ρ ranging from 0 to 1. A long recurrence interval was selected for this sensitivity analysis because it will tend to magnify the differences.

Figure 5-10 shows the variation of normalized stream concentrations (normalized with the concentration for the case of zero correlation) as a function of the correlation coefficient. The effect of flow correlation is substantial for large streams with a ratio of mean stream flow to mean runoff greater than 25, and minor for small streams with a corresponding ratio of less than unity. Flow variability, reflected by the ratio of stream to runoff coefficients of variation, has relatively minor effect, as illustrated by the width of the bands for each of the flow ratios.

Although Figure 5-10, indicates that flow correlations could significantly influence the magnitude of 10-year MRI stream concentrations, these results should be interpreted by considering the following. Streams with small flows relative to urban storm flows would reflect headwater segments and would most likely have high correlations. However, such streams are relatively insensitive to flow correlation effects on stream concentrations. On the other hand, the large streams which show the most dramatic influence are likely to represent situations in which flow correlations are small or nonexistent.

When these considerations are taken into account, the general assessment is that flow correlation is unlikely to influence the 10-year MRI stream concentrations by more than about 10 or 15%. Therefore, unless site-specific data are available to establish a local value for flow correlation, the assumption of a zero correlation is appropriate. Projections may be slightly conservative, but in most cases by no more than 10 percent or so.

5.4.3 Effect of Upstream Concentrations

The influence of upstream concentrations of a pollutant on water quality projections, and on the mitigation of impacts resulting from control measures is illustrated by the following three figures. The influence of significant upstream concentrations of a pollutant on quality below the stormwater discharge is shown by Figure 5-11. The solid curve applies for input conditions which are identical to those used for the example problem presented earlier in Chapter 3, where upstream concentration is assigned a value of zero. The dashed curve reflects the stream concentration distribution when the mean upstream (background) concentration is 50% of the mean runoff concentration. The same variability is assumed for the upstream concentrations as for the runoff concentrations.

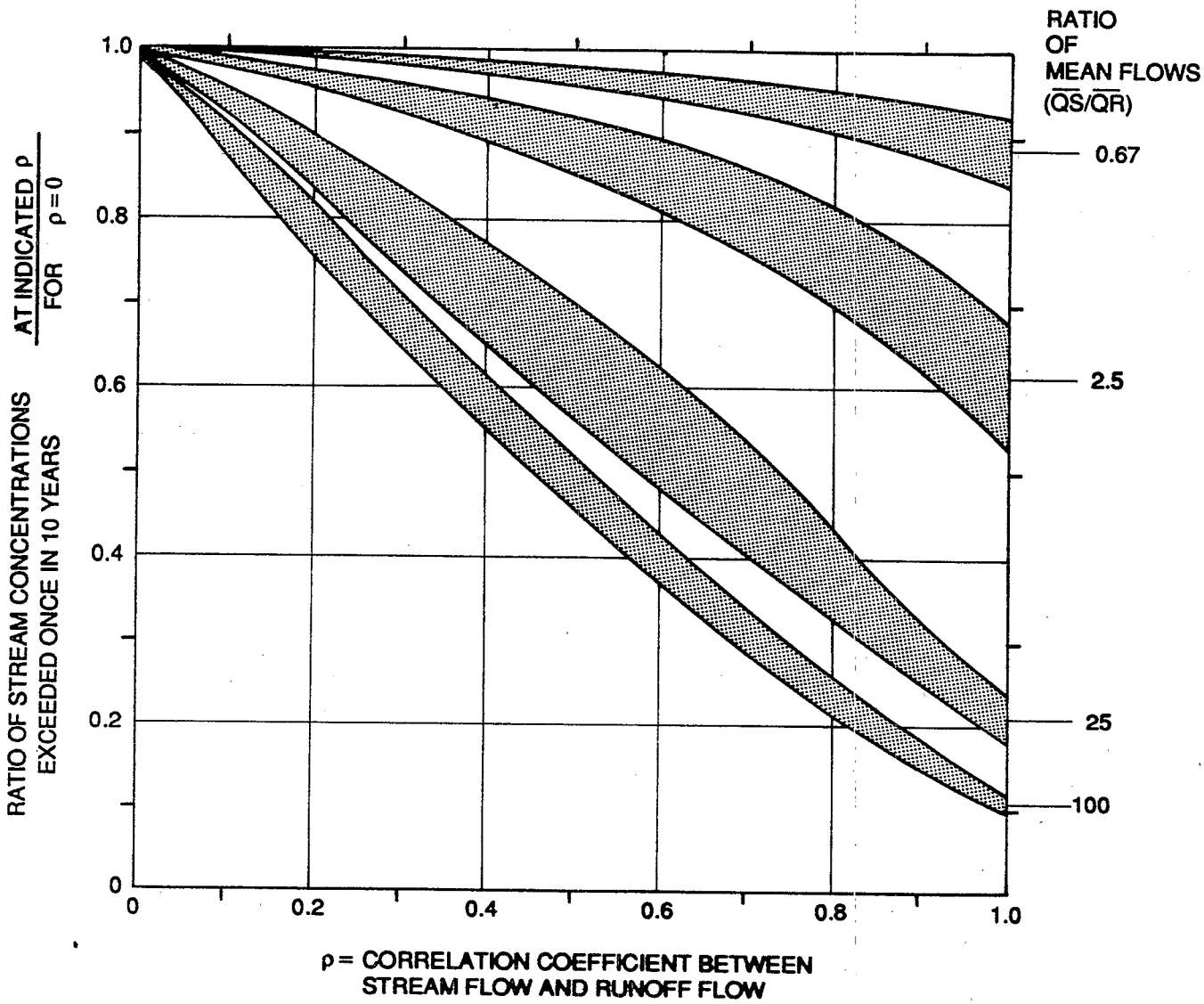


Figure 5-10. Effect of stream/runoff flow correlation

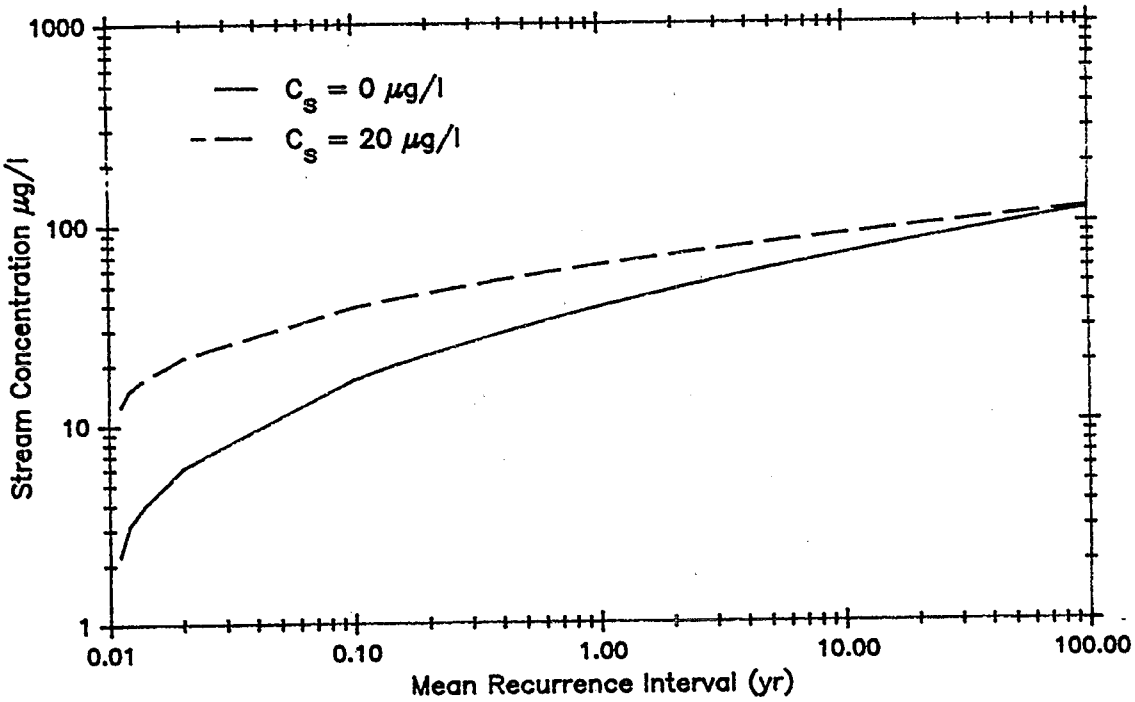


Figure 5–11. Effect of Upstream Concentration on Mean Recurrence Interval

The significance of this effect can be assessed by considering the degree of improvement in stream concentrations that result from a control action that produces a 50% reduction in the mean runoff concentration and no change in variability. Figure 5-12 illustrates the projected water quality impact of the control action for the case where upstream concentration is zero. The 50% reduction in urban runoff concentration would reduce the 3 year MRI concentration to 26 µg/1, and meet the desired objective. However, when the mean background concentration is close to the target concentration concentration, a 50% reduction in urban runoff produces little improvement in overall stream conditions as illustrated by Figure 5-13.

The practical significance of background concentrations (or those due to point sources or other upstream nonpoint sources) will vary with the particular combination of input conditions that characterize the local area. Stream and runoff flow relationships and variabilities will influence results, however, the factor of major importance will be background concentration levels relative to the target concentration and the runoff concentration.

5.5 SUMMARY AND DISCUSSION

The Derived Distribution Method has been shown to produce sufficiently accurate and usually conservative estimates of stream quality at recurrence intervals greater than a year, when compared with the results of the Direct Numerical Solution and the Monte Carlo Approach. The sensitivity analysis provides a basis for estimating the approximate magnitude of the degree to which 3-year recurrence concentrations are overestimated by the approximate method.

The magnitude and variability of upstream concentrations of pollutants have not been analyzed in detail. However, as the sensitivity analysis indicates, they can have a significant influence on control decisions. Where the upstream concentration of a pollutant is comparable to the target concentration, substantial reductions in urban runoff discharages may result in only marginal improvements in overall water quality and use impairment.

A high positive correlation between runoff flow and streamflow could result in significantly lower in-stream concentrations. This implies that runoff discharges enter when stream flows are in the higher range of possible values. This results in a high dilution and lower in-stream concentrations. There is, at present, insufficient information available to estimate typical values for this correlation. However, the data (rainfall and streamflow) required to estimate this parameter are available and it would be useful to analyze them. In the absence of this information, the assumption that the flows are uncorrelated provides conservative projections. The sensitivity analysis performed suggests that the assumption of no correlation probably overestimates 3-year stream concentrations by no more than 10-15 percent. This assumption is probably approximately correct for situations with large drainage area ratios (DARs), i.e., where the headwaters of the receiving stream are a considerable distance upstream of the urban area. In these larger streams, flow correlations would have the greatest effect. A more critical situation can be expected to occur in smaller urban streams with low DARs, where high correlations would be expected. However, in such cases runoff flows dominate conditions to the extent that the degree of flow correlation has little effect on stream concentrations.

The two analytical portions of the probabilistic dilution model presented embody the assumption that flows and concentrations are lognormal and independent. A large body of

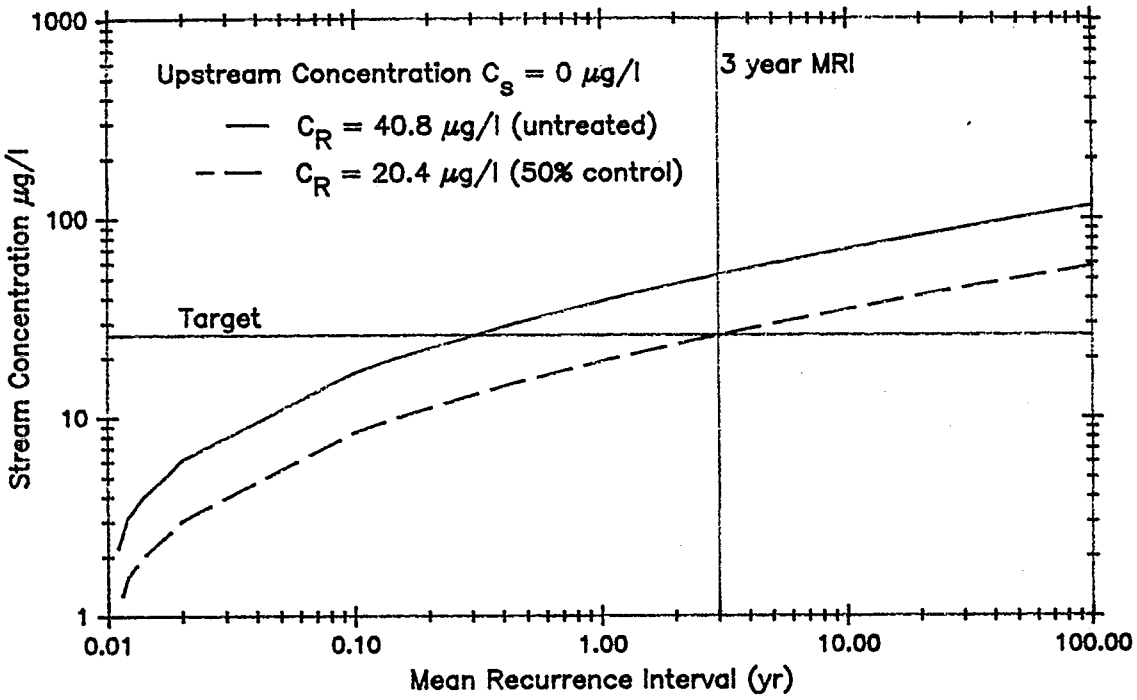


Figure 5–12. Effect of Control with No Background Contribution

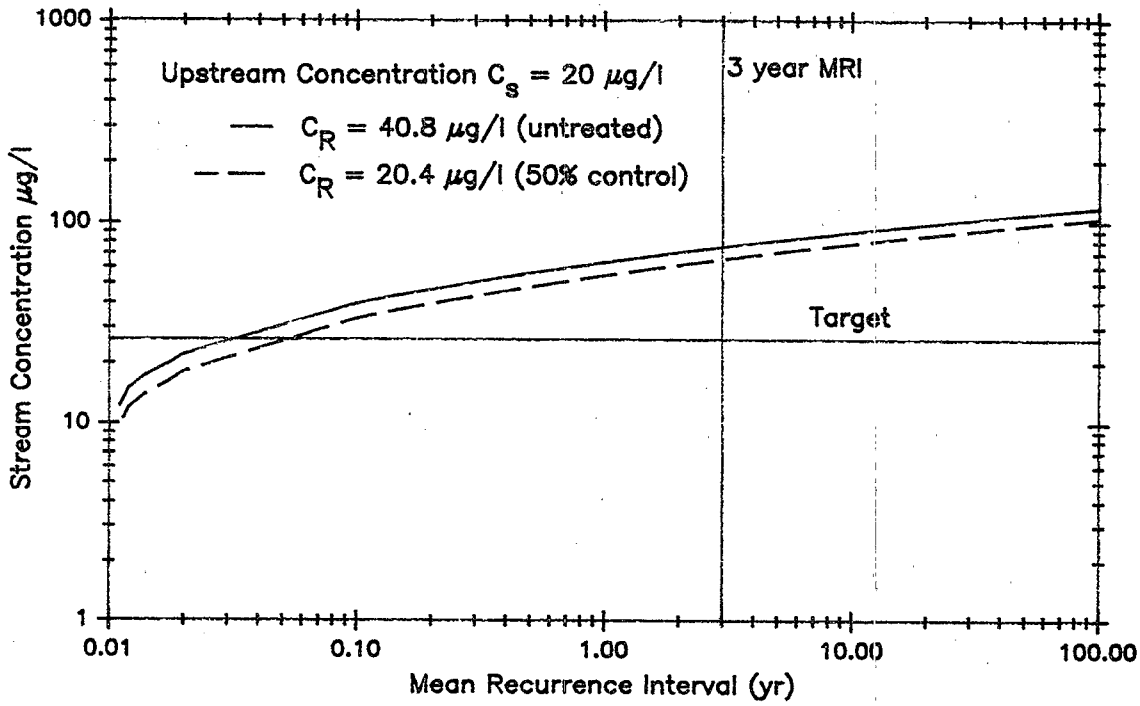


Figure 5–13. Effect of Control with High Background Contribution

data gives strong support to the appropriateness of this assumption in the case of urban runoff. For upstream concentrations and flows, nonzero correlations may be more likely but the examination of a sufficiently large number of data sets will be required before the significance of this assumption can be determined. Limited data review suggests that an assumption of lognormality is probably reasonable for upstream concentrations in most cases, but further investigation in this area would be useful. If, as in the case of the example analysis, upstream concentrations are set at zero to examine the effect of only the runoff discharge, then the issue of concentration lognormality and correlation with streamflow becomes moot.

The main issue relating to the assumption of lognormality of the inputs is in the case of the stream flow. Because of the presence of base flow in many streams, the current 2-parameter log normal distribution will tend to underestimate the lowest flows, and thus to overstate the magnitude of infrequently occurring stream concentrations. Some preliminary screening analyses on a few stream records suggest that this could be resolved by employing a 3-parameter log normal distribution, in which a minimum flow is also specified. However, with the current analysis, it should be recognized that output from the computation tends to produce conservative estimates for cases where the base flow component is significant.

The analytical methodology could be adapted to accept input parameters that are other than log normal. However, the computational structure is considerably simplified when all distributions are log normal, and the analysis is simple and easy to apply. In cases where input parameters are known to have distributions other than lognormal, the Monte Carlo technique may be more appropriate to use.

The basic approach and the specific procedures presented are useful for evaluating the significance of intermittent, variable pollutant discharges, and thus for providing guidance for rational and cost-effective water quality management decisions. Note that continuous point source discharges from publically owned treatment works (POTW) or industrial facilities also produce variable receiving water quality responses, and can be analyzed by essentially the same probabilistic methodology. The use of a probabilistic approach can provide an opportunity for evaluating both point and nonpoint source pollutant discharges using a similar framework, and hence evaluating point/nonpoint tradeoff control options on the basis of the effect on receiving water quality.

CHAPTER 6.0

USER'S GUIDE

This chapter provides guidance for using the Probabilistic Dilution Model (PDM) to estimate the in-stream concentrations and recurrence intervals. The theoretical basis of the model is described in detail in Chapter 3. A discussion of model application is given in Chapters 4 and 5.

This chapter is organized into five sections. Section 6.1 discusses the model acquisition and installation procedure. Section 6.2 presents the overall structure of the model. Common blocks and parameters used by the model are described in Section 6.3. A discussion of the batch input data file and interactive input is given in Section 6.4. An example application of the model is presented in Section 6.5. A list of the subroutines and the variables used in the model are included in Appendices C and D, respectively.

6.1 MODEL ACQUISITION AND INSTALLATION PROCEDURES

6.1.1. Model Acquisition

The PDM model along with the example input and output can be obtained by contacting:

Mr. Dennis Athayde
Nonpoint Source Branch
Criteria and Standards Division
U.S. E.P.A.
Waterside Mall, WH-554
Washington, D.C. 20460

The code is available on either 360K or 1.2M byte 5-1/4" - inch floppies for IBM PC-compatible machines, using DOS.

6.1.2. Model Installation

The source code for the model is contained in a file named PDM.FOR. This file has been compiled using the Ryan-McFarland FORTRAN compiler, and the resulting executable code, PDM.EXE, is also included. This executable version of the model requires a co-processor, a 8087, 80287 or 80387. If one is unavailable, the program needs to be recompiled using a compiler compatible with the computer running PDM. It is recommended that a backup copy of the original floppy be made before running the model. Consult the DOS manual for instructions on copying disks. If the model is run from a hard disk, then a directory should be created on the hard disk using the MKDIR (MD) command (consult your DOS manual for details on using this command). Once the directory has been created copy all of the files from the floppy into the directory.

The interactive portion of the model requires that the device driver ANSI.SYS,included with DOS, be installed. This device driver is installed by including the command:

DEVICE=ANSI.SYS

in the CONFIG.SYS file. After installing the device driver the computer needs to be rebooted. (consult you DOS manual for instructions on altering the CONFIG.SYS file).

To run the model, put the backup copy of the model disk in Drive A or access the directory on the hard disk containing the model and type PDM. The model will prompt you for the information it needs to run. If the model is to be run in batch mode, make sure that the input file is in the same directory as the model (if using a hard disk) or on the floppy in Drive A (if using a floppy disk).

If changes are made to any of the parameter statements in the code described in Section 6.3, the entire code must be recompiled and linked. The code has been developed using a Ryan McFarland version 2.0 compiler and a PLINK86 linker. The code requires the /b option when compliling. When compiling, all FORTRAN and INCLUDE files (described in Section 6.3) must be in the same directory and accessible to the user's compiler.

6.1.3. Testing the Model

Input and output files for the example problem discussed in Chapter 4 have been included with the model for both the analytic and Monte Carlo options. This example problem should be run and the results compared to the output file supplied with the model to ensure proper installation of the model.

6.2. MODEL STRUCTURE

The structure of the model is shown in Figure 6-1. The structure of the analytic model (shown Figure 6-2.) follows the steps presented in Section 3.3 for computing the exceedance probabilities. The model is divided into three components: input (either batch or interactive), output and computational subroutines. Subroutine descriptions are provided in Appendix C, and FORTRAN variables are defined in Appendix D. Subroutines used for low level control as part of the interactive preprocessor (e.g., calculating the length of strings, calculating screen coordinates, etc.) are not discussed or included in Appendix C or D.

6.3. COMMON BLOCKS AND PARAMETER STATEMENTS

Variables are passed between subroutines in the model through the use of common blocks and as arguments to subroutine calls. The computational portion of the analytic model uses a total of six common blocks, each containing a related set of parameters. These common blocks are contained in a file named PDM.CMN which is accessed by the code through the use of INCLUDE statements located at the beginning of a subroutine. The common blocks and a description of the parameters they contain are given below.

Common Block	Description
GENERAL	General run data
QSCMN	Stream flow and concentration data input by the user
QRCMN	Runoff flow and concentration data input by the user
CONC	Concentration data used by the numerical integration routines

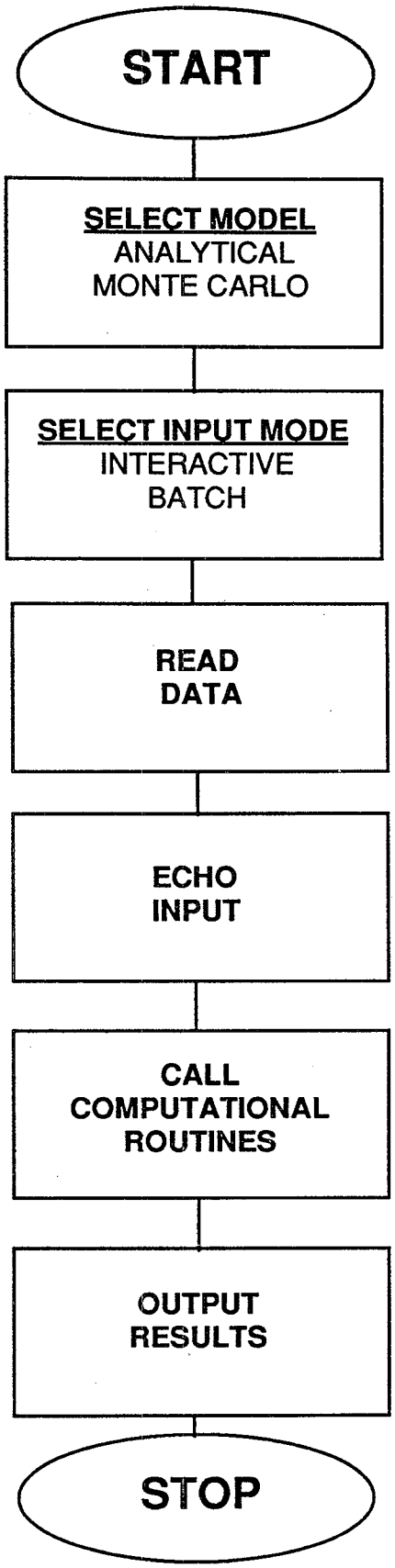


Figure 6-1 Flow Chart of the PDM Model

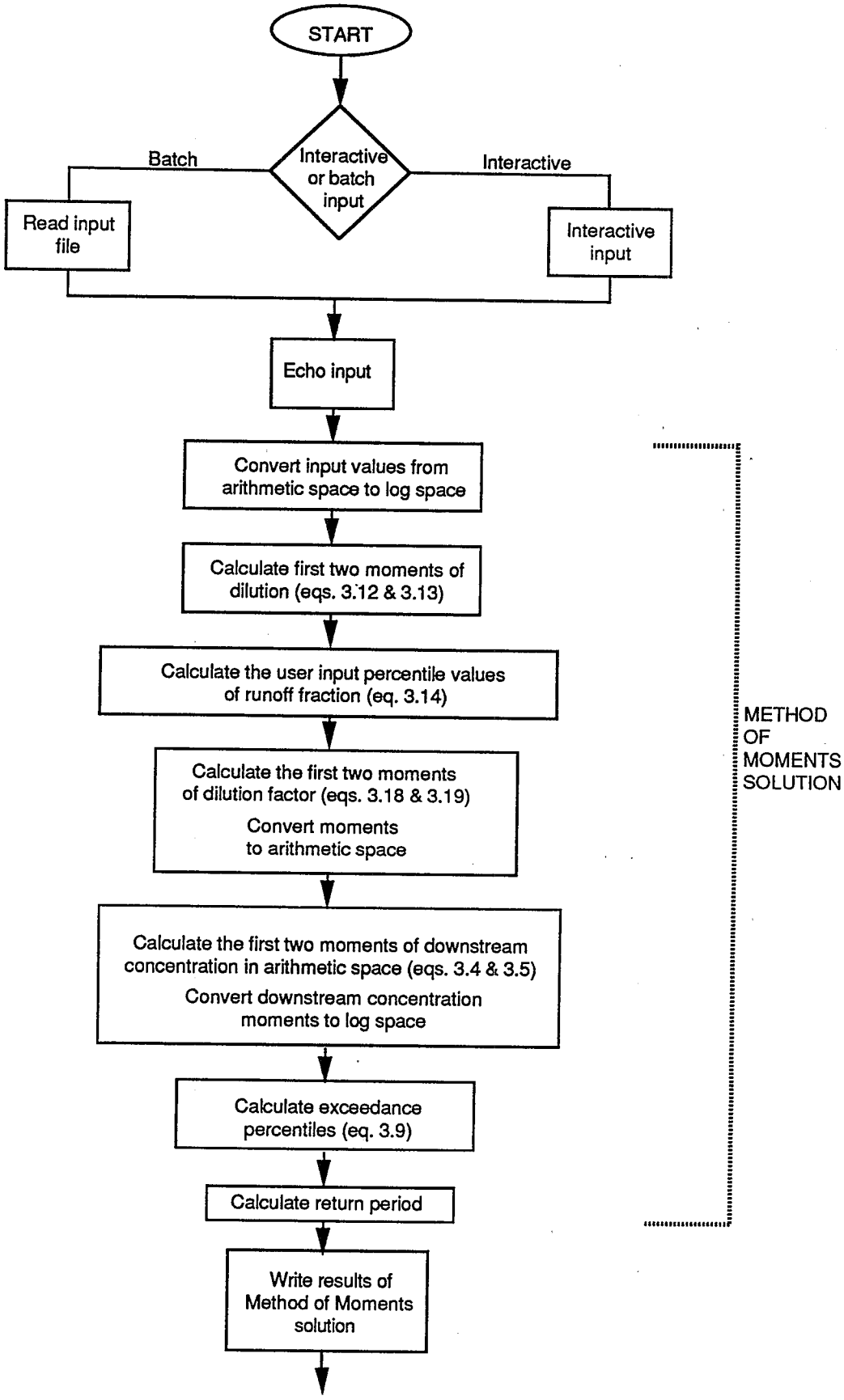


FIGURE 6-2 FLOWCHART OF THE ANALYTICAL PDM MODEL

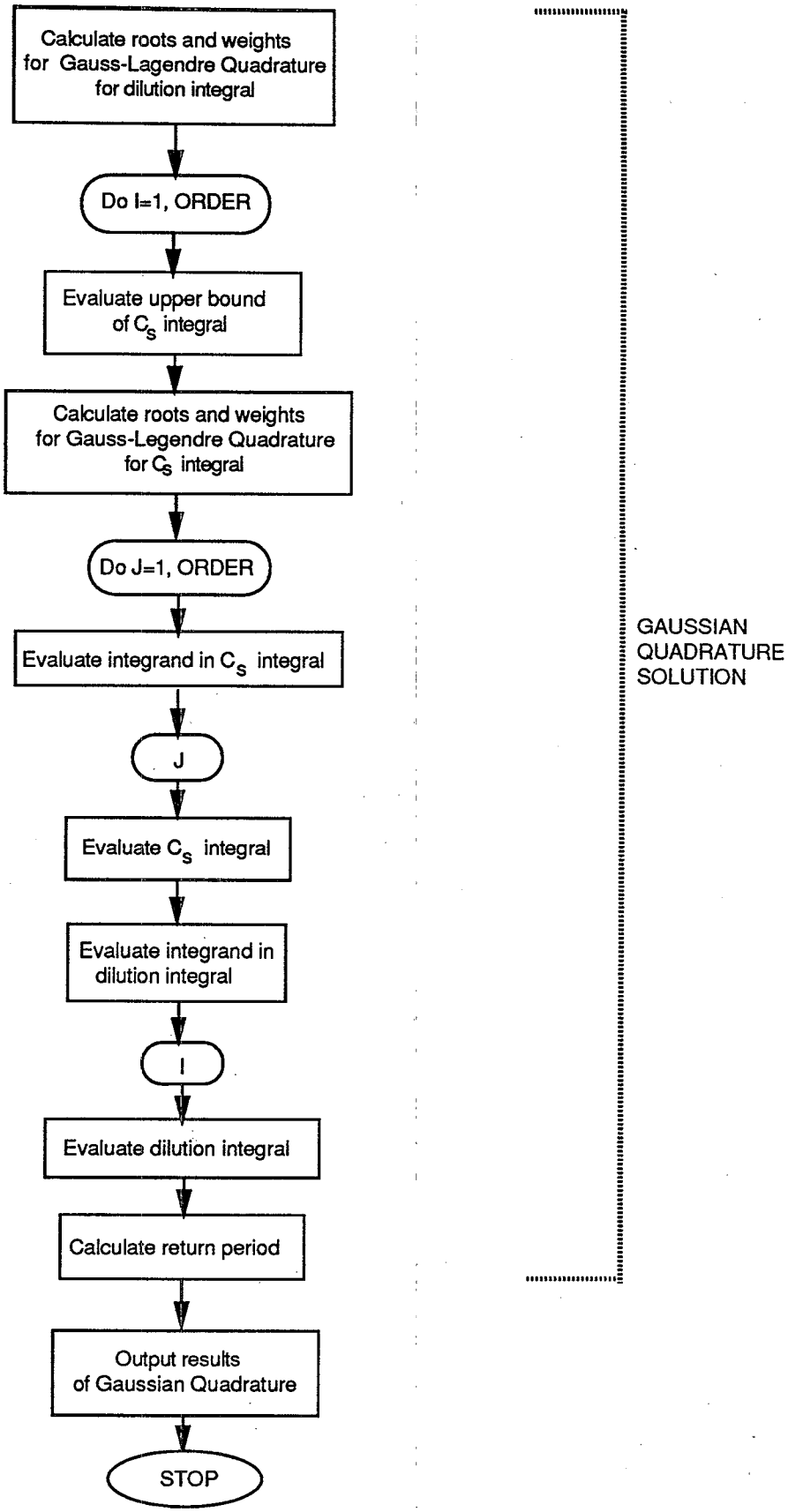


FIGURE 6-2 FLOWCHART OF THE ANALYTICAL PDM MODEL (concluded)

QSLCMN	Statistics of stream flow and concentration calculated by the model
QRLCMN	Statistics of runoff flow and concentration calculated by the model

In addition to the common blocks, the file PDM.CMN contains one parameter statement defining the value of two parameters. These parameters are used to dimension arrays required for calculating percentiles and return periods. The parameters and their assigned values are defined as follows:

Parameter Name	File	Value	Description
ECXMAX	PDM.CMN	10	Maximum number of sets of exceedances and return periods that can be calculated using the numerical integration routines
NPERCT	PDM.CMN	10	Maximum number of exceedances and return periods that can be calculated using the method of moments

These values are assigned in the file PDM.CMN and can be changed if necessary. Note the program has to be recompiled and linked if the values are changed.

The interactive preprocessor uses two additional common blocks and four parameter statements. The common blocks and the include files containing them are listed below.

Common Block	File	Description
SCREEN	SCREEN.CMN	Number of lines in program title
LINEST	SCREEN.CMN	ASCII codes used to define horizontal and vertical lines

Parameter statements used by the interactive preprocessor and the INCLUDE files containing them follow.

Parameter Name	File	Parameter Value	Description
HTYPE	SCREEN.PAR	1	Indicates that single lines are to be used for horizontal lines
VTYPE	SCREEN.PAR	2	Indicates that double lines are to be used for vertical lines
KIN	USRFILES	5	Unit number used to read from the keyboard

KOUT	USRFILES	6	Unit number used to write to the screen
CHR1WT	USRFILES	FALSE	Used to indicate if first character in output formats should be a blank or not
FXISTS	CMPLRERR	2013	Compiler error number indicating that a file exists

6.4. MODEL INPUT

Two sets of data are used by the PDM model, data for the analytical approaches and data for the Monte Carlo Approach. For each case data can be input in one of two ways (i) read from a file or (ii) entered interactively from the keyboard.

6.4.1 Analytical Model Data Requirements

Although the data requirements for the model are minimal, for ease of presentation it has been divided into three groups: general run data, flow and concentration data, and exceedence data. The formats necessary when using the batch input file are discussed below and shown in Table 6-1. When the model is run in the interactive mode, the model will prompt the user for input and no further guidance is required.

In addition to the data required by the model, comment lines can be included in a data file. Comments are identified by the presence of three asterisks (***) as the first three nonblank characters on a line. This is illustrated in the example input file presented in Table 6-2. Blank lines in a data file are treated as comment lines. Any system of units may be used for flow and concentration, but units must be consistent (i.e., all flows in cfs, all concentrations in mg/l).

The general run data consist of five parameters: the correlation between streamflow and runoff flow, the two quantiles used to estimate the mean and standard deviation of the dilution parameter, the number of storms per year, and the order of the gaussian quadrature approximation. The formats for entering these data are presented in Table 6-1. Note that the order of the gaussian quadrature approximation (ORDER) is restricted to 4, 5, 6, 10, 15, 20, 60, 104, and 256.

The exceedance probabilities (used in the Derived Distribution Appproach) for which the concentrations are to be calculated are input next. The format is presented in Table 6-1. A maximum of NPERCT (defined in the INCLUDE file PDM.CMN) values can be read in.

Flow and concentration data are input next using the formats presented in Table 6-1. Data associated with the stream are input first followed by data for the runoff.

The Direct Numerical Method calculates the exceedance probability for a given concentration. The concentrations used in calculating the exceedance probabilities are input next. The model reads in the concentrations in groups of three, a minimum concentration, a maximum concentration and an increment of concentration. The model calculates an exceedance probability at the minimum concentration and at the user-defined increments between the minimum and maximum concentrations. (For example, if the user read in 1, 100, 10 the model would calculate the exceedance probability for the concentrations 1, 11, 21, 31, 41, 51, 61, 71, 81 and 91.) In all ECXMAX (defined in the INCLUDE file PDM.CMN) sets can be read in. Table 6-1 presents the formats for input.

TABLE 6-1

INPUT FORMATS FOR ANALYTIC PDM MODEL

LINE	PARAMETER	FORMAT
1	RHO, ALPHA1, ALPHA 2, NST, ORDER	3F10.0, 2I10
2	PROBLS(I)	8F10.0
3	MQS, CVQS, MCS, CVCS	4F10.0
4	MQR, CVQR, MCR, CVCR	4F10.0
5	LCONC(I), HCONC(I), ICONC(I)	3F10.0
6	END	

VARIABLES	UNITS	DEFINITION
RHO	---	Correlation between stream flow and runoff
ALPHA1, ALPHA2	---	Z-score of quantiles used to estimate mean and standard deviation of dilution of the lognormal distribution
NST	---	Number of storms per year
ORDER	---	Order of gaussian quadrature approximation, must be one of the following: 4, 5, 6, 10, 15, 20, 60, 104, 256
PROBLS(I)	fraction	Exceedance probability for which concentration and return period are to be calculated using the Method of Moments
MQS	L^3/T	Mean streamflow
CVQS	---	Coefficient of variation of streamflow
MCS	M/L^3	Mean upstream concentration
CVCS	---	Coefficient of variation of upstream concentration
MQR	L^3/T	Mean runoff flow
CVQR	---	Coefficient of variation of runoff flow
MCR	M/L^3	Mean runoff concentration
CVCR	---	Coefficient of variation of runoff concentration
LCONC	M/L^3	Lowest concentration for which return period is desired in the Gaussian Quadrature Approximation

TABLE 6-1

INPUT FORMATS FOR ANALYTIC PDM MODEL (concluded)

VARIABLES	UNITS	DEFINITION
HCONC	M/L^3	Highest concentration for which return period is desired in the Gaussian Quadrature Approximation
ICONC	M/L^3	Increment of concentration for which return period is defined in the Gaussian Quadrature Approximation
END	---	The word END marking the end of the input file

Table 6-2
Sample Input file PDM1.IN for Analytic PDM Models

```
****      TEST DATA SET FOR PDM MODEL
***       TEST DATA IS FROM USERS MANUAL
***
*** GENERAL DATA
***    RHO      ALPHA1    ALPHA2       NST        ORDER
       0.0      1.282      2.652       100         15
***
***              PROBABILITY DATA
  .9998      .9900      .9800      .9500      .9000      .8000      .7000      .6000
  0.200      0.100
***
***      STREAM FLOW AND CONCENTRATION DATA
***   MQS        CVQS        MCS       CVCS
     125.0       1.25      0.010        0.1
***
***      RUNOFF FLOW AND CONCENTRATION DATA
***   MQR        CVQR        MCR       CVCR
       12.       1.15       40.8       0.60
***
***  LCONC      HCONC      ICONC
       10.       100.        10.
       .5         10.        .5
END
```

The last line in the data file should contain the letters 'END'. This signifies the end of all data.

6.4.2 Monte Carlo Data Requirements

This section describes the format of the data for the Monte Carlo simulation approach. Information presented in this section is further illustrated by the sample input file shown in Table 6-3.

The Monte Carlo input file consists of data lines and two types of general utility lines. The first type of utility line is the coment line indicated by the presence of three asterisks ('***'), (also see Section 6.4.1), as the first nonblank characters in the line. These lines are ignored and are provided to allow the user to type in comments, table headings, and other information useful in making the input file more understandable. Comment lines may be inserted anywhere in the data set. The second type of utility line is the END line used to mark the end of specific data groups. These lines are indicated by the word 'END' in the first three columns of the input line, and should be used only where specified in the following discussion.

Monte Carlo input data are comprised of five data groups: (1) simulation control parameters, (2) input distribution parameters, (3) empirical distribution data, (4) output options, and (5) correlated variable input. Data are read sequentially starting with Data Group 1 and ending with Data Group 5. Specific formats for each Data Group are shown in Tables 6-4 through 6-8 and are discussed below.

Data Group 1 --Simulation Control Parameters - This data group consists of two lines of data describing simulation options. Formats for these data are shown in Table 6-4. The first line contains the (alphanumeric) title for the run and is used to label the output. The second line contains the number of Monte Carlo runs to be used in the simulation, and the confidence level (percent) for estimating the confidence bounds for the quantile to be estimated.

Data Group 2 -- Input Distribution Parameters - This data group, shown in Table 6-5, consists of one line of data for each model variable (CS, CR, QS, QR, CO). The first entry on each line is a label, used to identify the random variable. The first two characters in the label must correspond to the labels shown in Table 6-4. The next 18 characters are user specified. The remaining data on these lines consist of parameters of the frequency distribution for the selected variables. After a data line is provided for each desired random variable, an END card must be supplied to mark the end of this data group. If data for any variable is not included in the input file it will be taken as a constant with a value of zero. Note that by setting the distribution flag VAR (5) to 0, the user can specify a constant value for a model input. In this case, the mean value of the variable (VAR (1)) will be used in the simulations. This option allows the user to vary the variables to be randomly generated without extensive modification of the input file.

Data Group 3 -- Empirical Distribution Data - This data group, shown in Table 6-6 contains the piecewise linear descriptions of cumulative frequency distributions for empirically distributed variables specified in Data Group 2 (by assigning a value of 5 to VAR(5)). The first data line for each distribution contains the number of data pairs NDAT used to describe the cumulative distribution and the name of the variable. This is followed by NDAT data lines, each containing (1) a value of the variable and (2) the corresponding cumulative probability for the specified value. These data pair lines must be input in ascending order of cumulative probability. If no variables have empirical distributions, this data group is not necessary. No END card is required for this data group.

Table 6-3
Sample Output File PDM1MC.OUT

```
EXAMPLE 1 FOUND IN PDM MANUAL
***
***  CONTROL PARAMETERS
***
 2000
***
***  MONTE CARLO INPUTS
PARAMETER NAME          MEAN      STD DEV.     MIN       MAX      DIST TYPE  ***
QS                      125.      156.25       0.0    10000.           2
QR                      12.0       13.80       0.0     5000.           2
CS                      0.01       0.01        0.0       1.0           2
CR                      40.8       24.48       0.0     5000.           2
END
***  EMPIRICAL DISTRIBUTION
***  OUTPUT CONTROL
QS                  NO CDF                WRITE
QR                  NO CDF                WRITE
CS                  NO CDF                WRITE
CR                  NO CDF                WRITE
CO                  NO CDF                WRITE
END
***  CORRELATION
END
***  END INPUT
```

TABLE 6-4

FORMAT FOR READING SIMULATION CONTROL PARAMETERS (DATA GROUP 1)

LINE	PARAMETER	FORMAT
1	TITLE	A80
2	NRUNS, PALPH	I5, F10.0

VARIABLES	DEFINITION
TITLE	Run Title.
NRUNS	The number of Monte Carlo runs.
PALPH	The confidence level 1 - PALPH (percent) confidence bounds on quantiles. If a value is not specified, 90% confidence bounds are used.

TABLE 6-5

FORMAT FOR READING DISTRIBUTION PARAMETERS FOR INPUT VARIABLES (DATA GROUP 2)

LINE	PARAMETER	FORMAT
1*	PNAME, VAR(1), VAR(2), VAR(3), VAR(4), VAR(5)	A20 5F10.0

VARIABLES	DEFINITION
PNAME	A label of up to 20 characters identifying the variable to be randomly generated. The first two characters must be one of the following: CS - upstream concentration CR - runoff concentration QS - upstream flow QR - upstream flow
VAR(1)	The mean of the distribution (or the most frequent value for triangular distributions).
VAR(2)	The standard deviation of the distribution.
VAR(3)	The minimum value for the random variable.
VAR(4)	The maximum value for the random variable.
VAR(5)	A flag specifying the distribution type of the random variable: 0 = Constant 1 = Normal 2 = Log-normal 3 = Exponential 4 = Uniform 5 = Empirical distribution, to be supplied in Data Group 3, Table 6-6 6 = Triangular

* Data lines repeated for each input variable.
The end of this Data Group is marked by an END card.

TABLE 6-6

FORMAT FOR READING EMPIRICAL DISTRIBUTION DATA (DATA GROUP 3)*

LINE	PARAMETER	FORMAT
1	NDAT, NAME	I5, A20
2a	DIST(I,1), DIST(I,2)	2F10.0

VARIABLES	DEFINITION
NDAT	Number of data pairs used to describe the piecewise linear cumulative distribution.
NAME	The name of the variable corresponding to the empirical distubution. (CS, CR, QS, QR)
DIST(I,1)	The value of the variable corresponding to the quantile for data pair I.
DIST(I,2)	The cumulative probability associated with the quantile DIST(I,1).

a This line is repeated NDAT times.

* Data Group repeated for each variable having an empirical distribution (VAR(5) = 5 in Data Group 2).

TABLE 6-7

FORMAT FOR READING OUTPUT OPTIONS (DATA GROUP 4)

LINE	PARAMETER	FORMAT
1*	(SNAME(I), I=13)	3A20

VARIABLES	DEFINITION
SNAME(1)	A label used to identify which variable is to be statistically summarized. First two characters must be from the list shown in Table 6-5 or a 'CO' can be used to indicate downstream concentration.
SNAME(2)	A flag that indicates if cumulative distributions should be plotted for the variable SNAME(1). This option is selected by putting "CDF" here. Putting anything else will result in the CDF not being plotted (i.e., NO COF).
SNAME(3)	A flag that indicates if values of the variable are to be written out for each Monte Carlo run (selected by putting "WRITE" here).

* Data lines repeated for each desired output variable. The end of this Data Group is marked by an END card.

TABLE 6-8

FORMAT FOR READING CORRELATED VARIABLE INPUT (DATA GROUP 5)

LINE	PARAMETER	FORMAT
1*	PNAME(1), PNAME(2), CORR(1,2)	2A20, F10.0

VARIABLES	DEFINITION
PNAME(1)	A 1- to 20-character label identifying the first correlated variable. Distribution type must be normal or log normal. First two characters must be CS, CR, QS, or QR.
PNAME(2)	A 1- to 20-character label identifying the second correlated variable. Distribution type must be normal or log normal. First two characters must be CS, CR, QS, QR or 'CO', the latter indicating downstream contamination.
CORR(1,2)	The value of the correlation coefficient for variables 1 and 2. Must be between -1 and 1.

* Data line repeated for each pair of correlated variables.
The end of this data group is marked by an END card.

Data Group 4 -- Output Options - The Monte Carlo simulation technique requires the application of the model for each combination of input parameters that is randomly generated. This process may be repeated several hundred times, resulting in several hundred values of model responses. The post-processor included in the model analyzes these concentration values to estimate (i) the mean, coefficient of variation, skewness coefficient, and kurtosis, (ii) the CDF of the output concentration, (iii) selected percentile values, and (iv) confidence intervals for the estimated percentile values. In addition, the post-processor includes a plotting subroutine that can be used to develop printer plots of the CDF. The results from these analysis are written to a file PDM.OUT.

This data group specifies the statistical output options for each variable to be written out. Data formats are shown in Table 6-7. This data group consists of one line for each output variable containing: (1) a character label up to 20 characters long identifying the output variable, (2) a flag indicating if a cumulative distribution should be plotted for this variable (selected by supplying the letters "CDF" here), and (3) a flag indicating if values of the variable random-generated are to written out for each Monte Carlo run (selected by supplying the word "WRITE" here). If "WRITE" option is chosen for any variable the results are written in file PDM.VAR. The labels used to identify variables must correspond to these specified in Table 6-5 through 6-7. A statistical summary table will be printed out for all variables selected in this data group. An END card is supplied to mark the end of this Data Group after a data line is supplied for each output variable.

Data Group 5 -- Correlated Input Variables - This data group, shown in Table 6-8, is used to indicate which of the input variables specified in Data Group 2 are correlated. Note that only variables with normal or log-normal distributions can be correlated. One line of data is provided for each pair of correlated variables. The first two entries on this line are labels identifying the two variables that are correlated. The third entry on these data lines is the value of the correlation coefficient. After a data line is supplied for each correlated pair of variables, an END card must be provided to mark the end of the data group. Note that an END card for this data group is always required even if no variables are correlated.

6.5. EXAMPLE APPLICATION

This section presents an example application that demonstrates the capabilities of the model. The input file necessary to run this example in the batch mode and the output file it produces are included on the floppy disk. Details of the example are discussed in Section 4.4 and Appendix B. The input file for this example is called PDM1.IN for the analytic mode. An example input file for the Monte Carlo option is also included called PDM1MC.IN.

Table 6-2 shows the input file for the analytical approach, Table 6-3 for the Monte Carlo Approach. The corresponding results are presented in Tables 6-9 and 6-10, respectively.

Table 6-9
Sample Output File PDM1.OUT

RECEIVING WATER IMPACT ANALYSIS

****

PROBABILITY DISTRIBUTION AND RETURN PERIOD

OF

RECEIVING WATER CONCENTRATIONS

Developed by

Woodward-Clyde Consultants
500 12th St., Oakland, CA. 94607
(415)893-3600

****

GENERAL MODEL RUN INFORMATION

Correlation between streamflow and runoff	.000
First Quantile for forced straight line agreemen	1.282
Second Quantile for forced straight line agreeme	2.652
Number of storms per year	100
Order of gaussian quadrature	15

CONCENTRATIONS ESTIMATED AT THE FOLLOWING EXCEEDANCE PROBABILITIES

1.000	.990	.980	.950	.900	.800	.700	.600
.200	.100						

	FLOW RATE		CONCENTRATION	
	MEAN	COEFFICIENT OF VARIATION	MEAN	COEFFICIENT OF VARIATION
Upstream of Discharge	125.00	1.25	.01	.10
Runoff	12.00	1.15	40.80	.60

Table 6-9 (continued)
Sample Output File PDM1.OUT

Stream concentrations percent exceedance and return period calculated for

Lowest	Highest	Increment
10.000	100.000	10.000
.500	10.000	.500

Downstream Concentration Statistics
Mean = 8.3 Standard Deviation = 7.8 Coef. of Var. = .93

ESTIMATES OF CONCENTRATIONS
USING THE METHOD OF MOMENTS APPROXIMATION

EXCEEDANCE PROBABILITY	CONCENTRATION	RETURN PERIOD YEARS
.000200	100.0873	50.007
.010000	38.3399	1.000
.020000	30.9056	.500
.050000	22.3661	.200
.100000	16.7790	.100
.200000	11.8460	.050
.300000	9.2162	.033
.400000	7.4381	.025
.800000	3.1307	.012
.900000	2.2103	.011

Table 6-9 (concluded)
Sample Output File PDM1.OUT

ESTIMATES OF EXCEEDENCE PROBABILITY
USING GAUSSIAN QUADRATURE SOLUTION

PROBABILITY	CONCENTRATION	RETURN PERIOD YEARS
.000181	100.0000	55.382
.000314	90.0000	31.875
.000563	80.0000	17.774
.001046	70.0000	9.557
.002031	60.0000	4.924
.004146	50.0000	2.412
.009007	40.0000	1.110
.021181	30.0000	.472
.055567	20.0000	.180
.175212	10.0000	.057
.175212	10.0000	.057
.187099	9.5000	.053
.200028	9.0000	.050
.214122	8.5000	.047
.229524	8.0000	.044
.246403	7.5000	.041
.264958	7.0000	.038
.285426	6.5000	.035
.308090	6.0000	.032
.333296	5.5000	.030
.361466	5.0000	.028
.393124	4.5000	.025
.428933	4.0000	.023
.469739	3.5000	.021
.516640	3.0000	.019
.571092	2.5000	.018
.635050	2.0000	.016
.711138	1.5000	.014
.802649	1.0000	.012
.911458	.5000	.011

Table 6-10
Sample Output File PDM1MC.OUT

```
------------- MONTE CARLO -------------

MONTE CARLO ANALYSIS OF PROBABILISTIC DILUTION MODEL

                DEVELOPED BY
                PHIL MINEART
                ATUL SALHOTRA

            WOODWARD-CLYDE CONSULTANTS
                 OAKLAND, CA

                    FOR

          U.S. ENVIRONMENTAL PROTECTION AGENCY
            ENVIRONMENTAL RESEARCH LABORATORY
                  ATHENS, GA  30613

                  JULY 1989

---------------------------------------

EXAMPLE 1 FOUND IN PDM MANUAL

  I N P U T   D A T A
  -------------------

  NUMBER OF MONTE CARLO RUNS          =  2000

  NUMBER OF MODEL INPUT PARAMETERS
   BEING VARIED                       =     4

  CONFIDENCE LEVEL FOR PERCENTILES    =  90.00 %
```

NOTATIONS

C_0 = downstream concentration in the river [mg/l]

C_{0a} = the downstream concentration exceeded with a probability (1 - a) [mg/l]

C_R = pollutant concentration in the runoff [mg/l]

c_R = natural logarithm of C_R

C_S = upstream concentration in the river [mg/l]

c_S = natural logarithm of C_S

d = natural logarithm of D

DP = mean rainfall event intensity [T]

DR = mean runoff event duration [T]

$f(c_s|d)$ = conditional density of c_s given d

$f(C_R, C_S, D)$ = the joint probability density function for C_R, C_S, and D

$f(c_R|c_s,d)$ = conditional density of c_R given c_s and d

f(d) = marginal probability density of d

MRI = the mean recursion interval [years]

Q_P = mean rainfall event intensity [L^3/T]

Q_R = mean runoff event flow rate [L^3/T]

Q_S = upstream flow rate in the river [m^3/s]

Rv = average runoff coefficient [dimensionless]

VR = mean runoff event volume [L^3]

VP = mean rainfall event volume [L^3]

z_a = the standard normal a quantile

$\eta(C_0)$ = coefficient of variation of C_0 in the arithmetic space

$\mu(x)$ = arithmetic mean of the variable x

$\mu_{ln}(C_0)$ = mean of C_0 in the logarithmic (normal) space

NOTATIONS (concluded)

$r\ \ln(Q_S,Q_R)$ = the cross-correlation coefficient between $\ln(Q_S)$ and $\ln(Q_R)$

$s(x)$ = standard deviation of the variable x

$s_{\ln}(C_0)$ = standard deviation of C_0 in the logarithmic (normal) space

ϕ = the runoff flow fraction or dilution [dimensionless]

Appendix A

MATHEMATICAL DETAILS OF THE NUMERICAL PROCEDURE

APPENDIX A1
DERIVATION OF $P_r\{\phi < \phi^*\}$

The runoff fraction is defined as:

$$\phi = \frac{Q_R}{Q_R + Q_S} = \frac{1}{1+D} = \frac{1}{1+e^d} \tag{A-1}$$

where: $D = Q_R / Q_S$ and $d = \ln(D)$, which is a normal random variable. Thus, the a-quantile, ϕ_a, is such that:

$$\begin{aligned} a &= P_r\{\phi < \phi_a\} = P_r\{\frac{1}{1+e^d} < \phi_a\} \\ &= P_r\{d > \ln(\frac{1}{\phi_a} - 1)\} \end{aligned} \tag{A-2}$$

since taking reciprocals inverts the inequality and taking logs of each side does not. Normalizing the probability results in:

$$a = P_r\{\frac{d - \mu(d)}{s(d)} > \frac{\ln(\frac{1}{\phi_a} - 1) - \mu(d)}{s(d)}\} \tag{A-3}$$

The left hand side of the inequality is standard normal, z_a. Solving inside the brackets for ϕ_a:

$$\ln(\frac{1}{\phi_a} - 1) = \mu(d) - z_a\, s(d) = \mu_l(D) - z_a s_l(D) \tag{A-4}$$

and:

$$\phi_a = \frac{1}{1 + \exp[\mu_l(D) - z_a s_l(D)} \tag{A-5}$$

where z_a is the standard normal a-quantile. The reason the derivation is straightforward is that ϕ is a monotonic function of d which is easily invertable.

APPENDIX A2

EVALUATION OF THE INTEGRALS DEFINING $P_r\{C_0 > C_0^*\}$

Equation 3.36 is stated below:

$$P_r\{C_0>C_0^*\} = \int_{d=-\bullet}^{\bullet} f(d)\ \{\int_{-\bullet}^{\ln C_{S1}} f(c_S|d)\ Q_R^*\ dc_S + Q_S^*\}\ dd$$

$$= \int_{-\bullet}^{\bullet} f(d) \int_{-\bullet}^{\ln C_{S1}} f(C_S|d)\ Q_R^*\ dC_S dd + \int_{-\bullet}^{\bullet} f(d) Q_S^*\ dd \qquad \text{(A-6)}$$

A number of methods have been tried for the numerical computation of these integrals. The most successful is based on an inverse probablility transformation which removes the probability density from the integral and results in a finite range of integration. If the integrals are written using the lognormal probability density function:

$$P_r\{C_0 > C_0^*\} = \int_0^{\bullet} f(D)\ \{\int_0^{C_{S1}} f(C_S|D)\ Q_R^*(C_S, D) + Q_S^*(D)\}\ dD \qquad \text{(A-7)}$$

where f is the lognormal probability density function. Define new variables of integration, y and x, by the relationships:

$$y = \int_0^{C_s} f(C_s|D)\ dC_s = P_s(C_s) \qquad \text{(A-8)}$$

$$x = \int_0^{D} f(D)\ dD = P_D(D) \qquad \text{(A-9)}$$

Then:

$$dy = f(C_s|D)\ dC_s \qquad \text{(A-10)}$$

$$dx = f(D)\ dD \qquad \text{(A-11)}$$

The limits of integration need to be transformed also.

For D = 0

$$\int_0^0 f(D)\ dD = 0 \qquad \text{(A-12)}$$

For D = •

$$\int_0^{\bullet} f(D)\ dD = 1 \qquad \text{(A-13)}$$

For $C_S = 0$

$$\int_0^0 Q_R\, f(C_S|D)\, dC_S = 0 \tag{A-14}$$

For $C_S = C_{S1}$

$$\int_0^{C_{S1}} f(C_S|D)\, dC_S = P_r\,(C_S \,£\, C_{S1}) = P_S\,(C_{S1}) \tag{A-15}$$

Expressions for C_S and D are also needed. These can be written as inverse functions of the form $P^{-1}(y) = C_S(y)$ and $P^{-1}(x) = D(x)$.

With these substitutions, equation A-7 can be rearranged to give:

$$P_r\{\, C_0 > C_0^*\} \int_0^1 \{\int_0^{P_S(C_{S1})} Q_R^* [C_S(y), D(x)]\}\, dy + Q_S^* [D(x)]\}\, dx \tag{A-16}$$

which can now be evaluated using standard guassian quadrature formulas. It should be noted that $Q_R{*}$ is now a function of $C_S(y)$ and D(X) and $Q_S{*}$ is now a function of D(x).

The probability, $P_S(\, C_{S1}\,)$, and the inverse functions, $C_S(y)$ and D(x), are computed from the standard normal cumulative distribution function:

$$P(z^*) = P_r\,\{\, z < z^*\,\} = 1 - Q^*\,(z^*) \tag{A-17}$$

and its inverse. That is , if $p = P(z^*)$, then the inverse function, P^{-1}, satisfies the relationship:

$$z^* = P^{-1}(p) \tag{A-18}$$

Thus:

$$P_S(C_{S1}) = P_r\,\{C_S \,£\, C_{S1}\} = 1 - P_r\,(C_S \geq C_{S1}) = 1 - Q_S^*\,[D(x)] \tag{A-19}$$

The inverse functions for $C_S(y)$ and D(x) can be found using equation A-18:

$$\frac{\ln C_S(y) - \mu\,(C_S|d)}{s(C_S|d)} = P^{-1}(y) \tag{A-20}$$

Therefore

$$C_S(y) = \exp\,[\mu(c_S|d) + P^{-1}(y)\; s(c_S|d)\,] \tag{A-21}$$

and similarly

$$D(x) = \exp\,[\mu(d) + s(d)\,P^{-1}(x)] \tag{A-22}$$

Gauss-Legendre Quadrature was used to integrate equation A-16. In order to use this method the limits of integration need to be in the interval [-1, 1]. The roots of the Legendre Polynomials can be transformed to the interval defined by the actual limits using the following formula:

$$x_i = \frac{b+a}{2} + \frac{b-a}{2}\,x_i \tag{A-23}$$

where a and b are the upper and lower limits of integration, and the x_i's are the roots of the Legendre Polynomials. For the D integral, the roots are tranformed using:

$$x_i = \frac{1}{2} + \frac{1}{2}\,x_i \tag{A-24}$$

and for the C_S integral the roots are transformed using

$$y_{ij} = \frac{1 - Q^*_{si}}{2}\,(1 + x_j) \tag{A-25}$$

With these transformations the integrands in equation A-16 can be evaluated at the Legendre Polynomial roots. The D integral integrand is:

$$Q^*_{Si} = Q^*\,\left\{\frac{\ln\,[C^*_T(1 + 1/D_i) - \mu\,(c_S|d_i)]}{s\,(c_S|d_i)}\right\} \tag{A-26}$$

where D_i is evaluated as:

$$D_i = D\,(x_i) = \exp\,[\mu(d) + s(d)\,P^{-1}(x_i)] \tag{A-27}$$

where the conditional moments are given in Table A1.

The C_S integral integrand is:

$$Q^*_R\,[C_S\,(y_{ij}), D(x_i)] =$$

$$Q^*_R\,\left\{\frac{\ln\,[C^*_T(1+D_i) - D_i\,C_s(y_{ij})] - \mu\,(c_R|c_S, d)}{s\,(c_R|c_S, d)}\right\} \tag{A-28}$$

with the conditional moments evaluated at $C_S = \ln\,(C_S(y_{ij}))$ and $d = \ln D_i$. The integrals can now be solved using standard Gaussian Quadrature.

Appendix B

EXAMPLE APPLICATION OF THE METHOD OF MOMENTS METHOD

FIGURE B-1

EXAMPLE APPLICATION OF THE METHOD OF MOMENTS METHOD

Input Data	Unit	Mean	Coefficient of Variation
Stream Flow (Q_S)	cfs	125	1.25
Runoff Flow (Q_R)	cfs	12	1.15
Runoff Conc.(C_R)	µg/l	40.8	0.60
Upstream Conc. (C_S)	µg/l	0	-
Correlation of Q_S and Q_R			
$r(Q_S, Q_R)$		0	-

STEP 1: Compute statistical parameters of inputs assuming lognormal distributions.

• For runoff concentration (C_R):

Arithmetic

Mean $\mu(C_R)$ = ...(given)..... = 40.80

Coef. Var. $\eta(C_R)$ = ...(given)..... = 0.60

Std. Dev. $s(C_R) = \mu(C_R) \cdot \eta(C_R) = (40.8) \cdot ((0.6)$ = 24.48

Median $(\widetilde{C_R}) = m(C_R)/\sqrt{1+\eta^2(C_R)} = 40.8/\sqrt{1+(0.6)^2} = 34.99$

Logarithmic

Log Mean $\mu_{ln}(C_R) = \ln(\widetilde{C_R}) = \ln(34.99) = 3.555$

Log Std. Dev. $s_{ln}(C_R) = \sqrt{\ln(1+\eta^2(C_R)} = \sqrt{\ln(1+(0.6)^2)} = 0.555$

• Repeat these computations for the other input parameters, and tabulate results for convenience in the following steps:

	Arithmetic				Logarithmic	
	Mean	Median	Std.Dev.	Coef.Var.	Mean	Std.Dev.
Runoff Conc.(C_R)	40.80	34.986	24.48	0.60	3.555	0.555
Runoff Flow (Q_R)	12.00	7.874	13.80	1.15	2.064	0.918
Streamflow (Q_S)	125.00	78.087	156.25	1.25	4.358	0.970

STEP 2:

(a) Compute standard deviation and the log mean of the flow ratio $Q_S/Q_R = D$:

$$s_{\ln(D)} = \sqrt{s^2_{\ln(Q_S)} + s^2_{\ln(Q_R)} - 2r\, s_{\ln(Q_S)} s_{\ln(Q_R)}}$$

The first two terms are taken from the table just prepared (and squared). Since, for this example, flows are uncorrelated ($r = 0$), the third term drops out.

$$s_{\ln}(D) = \sqrt{(0.970)^2 + (0.918)^2} = 1.336$$

$$\mu_{\ln}(D) = \mu_{\ln}(Q_S) - \mu_{\ln}(Q_R)$$
$$= 4.358 - 2.064 = 2.294$$

(b) Compute the 90th and 99.6th percentiles of the actual distribution of the dilution factor (ø) using the appropriate values of z for the selected percentiles:

$$z_{90} = 1.282 \qquad z_{.996} = 2.652$$

$$\phi_a = \frac{1}{1+\exp[\mu_{\ln}(D) - z_a s_{\ln}(D)]}$$

Substituting the appropriate values gives:

$$\phi_{90} = 0.358 \qquad \phi_{.996} = 0.777$$

(c) Compute log standard deviation and log mean of lognormal approximation of the distribution of the dilution factor (ø):

Log std.dev. $s_{\ln}(\phi) = \dfrac{\ln(\phi_{99.6}) - \ln(\phi_{90})}{z_{99.6} - z_{90}} = 0.565$

Log mean $m_{\ln}(\phi) = \ln(\phi_{99.6}) - z_{99.6}\, s_{\ln}(\phi) = -1.750$

(d) Compute arithmetic transforms using equations in Table 4-1, and tabulate for convenience:

	Arithmetic				Logarithmic	
	Mean	Median	Std.Dev.	Coef.Var.	Mean	Std.Dev.
Dilution Factor (ϕ)	0.2038	0.1738	0.1249	0.61	-1.749811	0.564562

STEP 3: Compute the statistical parameters of the resulting in-stream concentration (C_0).

(a) Compute the arithmetic mean concentration using previously tabulated values:

$$m(C_0) = (m(C_R) \cdot m(\phi)) + [\, m(C_S) \cdot (1 - m(\phi))\,]$$
$$= (40.80 \cdot 0.2038) + [\,0\,] = 8.32$$

(b) Compute the standard deviation:

$$s(C_0) = \sqrt{s^2(\phi)\,[\,m(C_R) - m(C_S)\,]^2 + s^2(C_R)\,[s^2(\phi) + \mu^2(\phi)] + s^2(C_S)\,[s^2(\phi) + (1 - m(\phi))^2]}$$

$$= 0.0156 \cdot (40.80 - 0)^2 + 599.27 \cdot (0.0156 + 0.0415) + [\,0\,]$$
$$= 7.76$$

(c) Compute and tabulate other statistical parameters of stream concentration for use in subsequent analyses:

	Arithmetic				Logarithmic	
	Mean	Median	Std.Dev.	Coef.Var.	Mean	Std.Dev.
Stream Conc. (C_0)	8.32	6.08	7.76	0.93	1.805	0.791

STEP 4:

Use the statistical parameters of stream concentration computed in the previous step to construct a graphical display (or table) summarizing: (a) the probability distribution and (b) the recurrence interval.

(a) Probability plot using log probability paper (see Figure 4-3)

- Plot the median ($\widetilde{C_0} = 6.08$) at the 50% plotting position.
- From a table of standard normal distributions (any text on statistics) select a set of probabilities and their corresponding z scores. For example :

Probability (a)	z_a
0.50 (50%)	0
0.159 (16%)	-1.000
0.841 (84%)	+1.000
0.050 (5%)	-1.645
0.950 (95%)	+1.645

- For the selected probability values, estimate the corresponding downgradient percentile concentration using Equation 3.9 shown below:

$$C_{0a} = \exp\,[\mu_{ln}(C_0) + z_a s_{ln}(C_0)]$$

	Prob.	z	C_0
Log mean = 1.805128	5 %	-1.645	1.65
Log sigma = 0.791337	16 %	-1.000	2.75
	84 %	+1.000	13.43
	95 %	+1.645	22.36

(b) Plot the Mean Recurrence Interval (see Figure 4-4).

- Rearrange Equation 3.9 (used in previous step) to:

$$z_a = \frac{\ln(C_{0_a}) - \mu_{\ln}(C_0)}{s_{\ln}(C_0)}$$

- Select a series of values for stream concentration (C_0), take the natural (base e) log, and enter in the equation with the log mean and log sigma (determined in Step 3), to compute the value of z_a corresponding with the selected concentration.

- From a standard normal table, determine the probability (a) associated with z_a. Note that standard tables are usually set up so that a reflects the probability equal to or less than.

- To compute the mean occurence interval (MRI) at which the selected concentration will be equalled or exceeded:

$$\text{exceedance probability} = 1 - a$$

and

$$\text{MRI} = \frac{1}{(\text{exceedence prob.}) \cdot (\text{no.storms / yr})} = \text{years}$$

- For example, given the stream concentration distribution from Step 3 and 100 storms per year:

	Exceedance Prob.		MRI
C_{0a}	z_a	(1-a)	(years)
100	3.538	0.000201	49.7
80	3.257	0.000564	17.7
60	2.893	0.001909	5.2
50	2.663	0.003876	2.6
40	2.381	0.008645	1.2
30	2.017	0.021852	0.5
20	1.505	0.066220	0.15

Figure B-2(a) presents the cumulative probability distribution of the downstream concentrations. Figure B-2(b) presents the average recurrence interval associated with specific stream concentrations that will be produced during storm runoff periods.

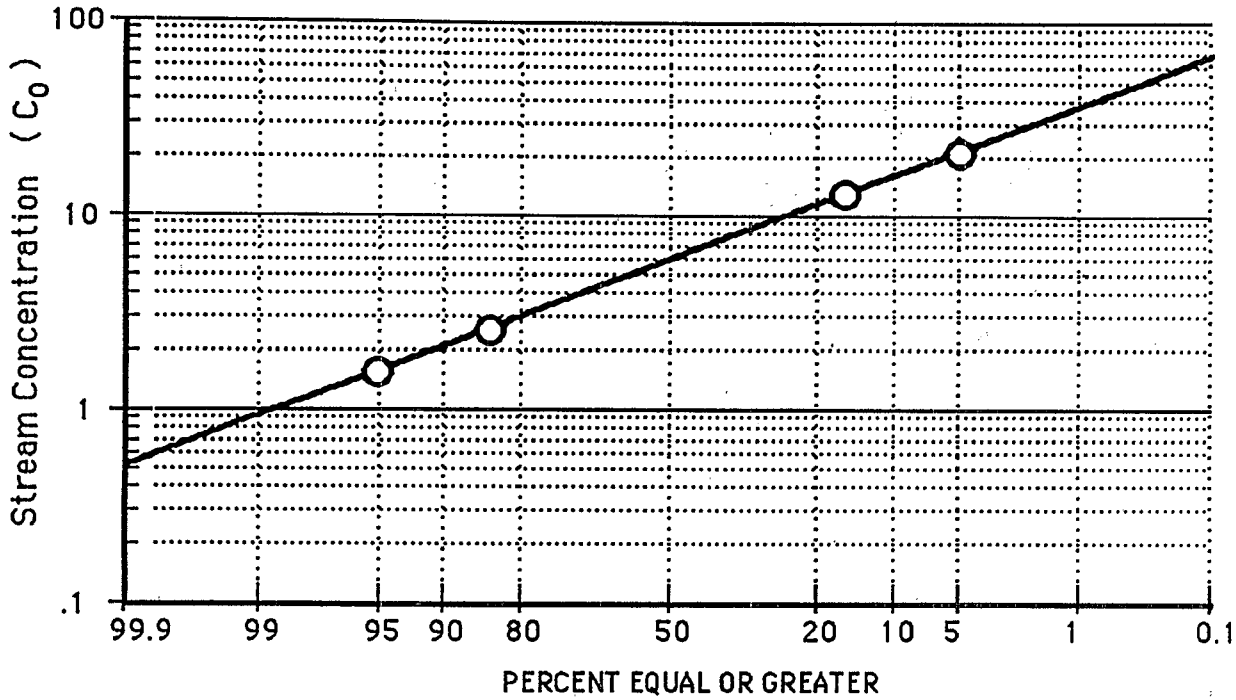


Figure B-2(a). Probability Distribution of Pollutant Concentrations During Storm Runoff Events

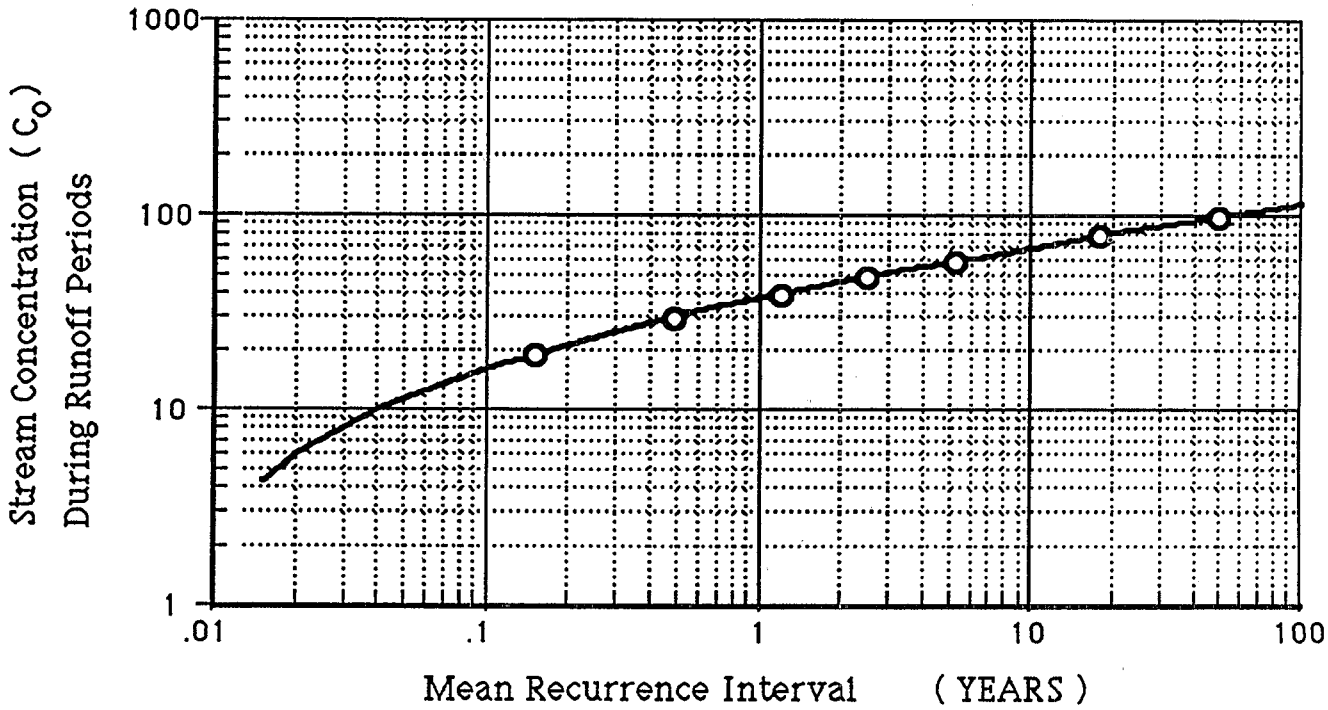


Figure B-2(b). Recurrence Intervals for Stream Concentrations

Appendix C

LIST OF SUBROUTINES INCLUDED IN THE PDM MODEL

TABLE C-1

LIST OF SUBROUTINES INCLUDED IN THE ANALYTIC PDM MODEL

SUBROUTINE	CALLED BY	DESCRIPTION
Input/Output Routines		
READF	DPDM	Reads data from a file in batch mode
COMRD	READF	Searches for data lines in input files
OPENF	READF, ECHO	Opens input and output files
ECHO	DPDM	Echos input data to output file
OUTPUT	DPDM	Writes results to output file
SCREEN	DPDM	Writes status screen to terminal
GENERL	DPDM	Interactively reads general run data
CONCIN	DPDM	Interactively reads flow and concentration data
CONCN2	DPDM	Interactively reads concentration data used by Gaussian Quadrature Method
EXCEED	DPDM	Interactively reads exceedance data
INTACT	GENERL, CONCIN, EXCEED	Interactively inputs a table of values
READR	INTACT	Reads input data from the screen
ERR	INTACT	Prints an error message to the screen when interactively inputting data
DPDM	MAIN	Main calling routine for analytical models
Computational Routines		
TRNLOG	DPDM	Transforms mean and standard deviation from the arithmetic space to log space

TABLE C-1

LIST OF SUBROUTINES INCLUDED IN THE ANALYTIC PDM MODEL (concluded)

SUBROUTINE	CALLED BY	DESCRIPTION
MOMENT	DPDM	Calculates exceedance probabilities using the method of moments
Function INVNRM	MOMENT	Approximates the inverse of the normal (gaussian) probability distribution function
NUMSUL	DPDM	Calculates the exceedance probability using direct numerical interaction of joint probability function using Gaussian Quadrature
GAUSLG	NUMSOL	Integrates an arbitrary real function as a single variable over a finite interval using Gaussian-Legendre Quadrature
DGAUSS	GAUSLG	Computer roots and weight fractures for the Gauss-Legendre Quadrature. Transforms the integrate interval to [-1,1]
Function NRMPRB		Approximates the normal probability function
Function QSTARR	GAUSLG	Function QR (see Chapter 3)

TABLE C-2

LIST OF SUBROUTINES INCLUDED IN THE MONTE CARLO PDM MODEL

Subroutine	Called By	Description
BMPCHR	NAMFIX	Capitalizes character variables.
CENTER	OPENF	Centers a character variable.
COMRD	READM	Reads and checks if input lines contain comments, data, or end of file.
CONOUT	MAIN	Reads output options in Monte Carlo mode.
DECOMP	INITMC	Decomposes correlation matrix into coefficient matrix required to generate correlated random numbers.
EMPCAL	RANDOM	Generates empirically distributed random numbers by interpolating linearly from an input empirical distribution.
ERR	VARBLE, CONOUT	Prints error message to screen.
ERRCHK	COMRD	Writes error messages when unexpected end-of-file is encountered in input data file.
EXPRN	RANDOM	Generates exponentially-distributed random numbers.
FRQPLT	OUTFOR	Writes plots of cumulative frequency distributions.
FRQTAB	OUTFOR	Writes tabulated cumulative frequency distributions.
INCOR	MAIN	Reads in correlative matrix.
INIT	MAIN	Initializes Monte Carlo variables.
INITMC	MAIN	Initializes statistical summation arrays, reorganizes Monte-Carlo input arrays to account for constant variables, and performs other miscellaneous Monte-Carlo intializations.
INTRO	MAIN	Writes initial screen.

TABLE C-2 (Continued)

LIST OF SUBROUTINES INCLUDED IN THE MONTE CARLO PDM MODEL

Subroutine	Called By	Description
LFTJUS	NAMFIX	Left-justifies character variables (i.e., removes blanks from the left-side of character strings).
LNGSTR	NAMFIX	Finds the length of character variables (i.e., the number of non-blank characters).
MCECHO	READM	Writes Monte-Carlo input data to the output file.
MTPV	RANDOM	Multiplies a vector of uncorrelated variables by a coefficient matrix to reform a vector of correlated variables.
NAMFIX	READM	Left-justifies and capitalizes character variables.
NMB	RANDOM	Generates normally-distributed random numbers.
OJTFOR	OUTPUT	Writes frequency tables and plots.
OUTPUT	MAIN	Writes out statistical summaries of Monte-Carlo runs to the output file.
PDM	MAIN	Monte Carlo computational routine.
QUES	MAIN	Asks user whether input in batch or interactive.
QUESST	OPENF	Asks user whether to overwrite a file.
QUES2	MAIN	Asks user which model to run.
RANDOM	MAIN	Generates a vector of random numbers from specified distributions.
READA	CONOUT	Reads character variable from screen.
READM	MAIN	Reads Monte-Carlo input data from a user-specified input file unit number.
READR	VARBLE	Reads real number from screen.

TABLE C-2 (Concluded)

LIST OF SUBROUTINES INCLUDED IN THE MONTE CARLO PDM MODEL

Subroutine	Called By	Description
REMP	MAIN	Reads empirical distributions for the Monte Carlo model in interactive mode.
STATIS	MAIN	Performs summations required to compute statistical moments for random model inputs and model outputs over all Monte-Carlo runs.
STOUT	OUTPUT	Computes statistical moments (mean, standard deviation, skewness, kurtosis, correlations, minimum and maximum) from summations computed by STATIS. Statistics are then written out to the Monte-Carlo output file.
UNIF	RANDOM,NMB, EXPRN	Generates uniform random numbers ranging between 0 and 1.
VARBLE	MAIN	Reads in Monte Carlo distribution data while in interactive mode.

Appendix D

LIST OF VARIABLES INCLUDED IN THE PDM MODEL

TABLE D-1
LIST OF VARIABLES INCLUDED IN THE PDM ANALYTIC MODEL

Variable Name	Units	Type*	Description	Subroutine /Function /Include File	Common	Code
ACVDIL	---	Scalar(R)	Coefficient of variation of dilution in arthmetic space	Moment		
ALPHA	---	Scalar(R)	Quantile used to estimate new and standard deviation of dilution	Pdm.cmn	General	
AMNDIL	---	Scalar(R)	Arithmetic mean of dilution	Moment		
ANS	---	Scalar(C)	Temporary storage variable for screen input	Readr		
ASDDIL	---	Scalar(R)	Standard deviation of dilution in arithmetic space	Moment		
BEEP	---	Scalar(C)	ASCII code for beep	Err		
BLANK	---	Scalar(C)	Blank Character	Comrd		
BRACT	---	Scalar(C)	ASCII code for square left bracket '['	Intact		
C0,C1,C2	---	Scalar(R)	Coefficients of approximation for inverse normal function	Invnrm		
COL	---	Scalar(I)	Column on screen used in interactive preprocessor	Openf, Intact		
COM	---	Scalar(R)	Downstream concentration	Main, Moment Output		O, I
CVCOM	---	Scalar(R)	Coefficient of variation for downstream concentration	Moment		
CVCR	---	Scalar(R)	Coefficient of variation of runoff concentration	Pdm.cmn	Qrcmn	
CVCS	---	Scalar(R)	Coefficient of variation of upstream concentration	Pdm.cmn	Qscmn	
CVQR	---	Scalar(R)	Coefficient of variation of runoff	Pdm.cmn	Qrcmn	
CVQS	---	Scalar(R)	Coefficient of variation of streamflow	Pdm.cmn	Qscmn	
D1,D2,D3, D4,D5,D6	---	Scalar(DP)	Coefficients of approximation to normal distribution	Nrmprb		
DATAEND	---	Scalar(L)	Logical variable used to indicate end of data file	Readf		O
CR	mg/L	Scalar (R)	Runoff concentration	PDM		

TABLE D-1
LIST OF VARIABLES INCLUDED IN THE PDM ANALYTIC MODEL (continued)

Variable Name	Units	Type*	Description	Subroutine /Function /Include File	Common	Code
DEFAULT	---	Scalar(C)	Used for writing default values to screen in interactive preprocessor	Intact		
DIR	---	Scalar(C)	Flag used for interactive input	Readr		M
DONE	---	Scalar(L)	Flag used to indicate end of interactive input	Intact		
ECXMAX	---	Scalar(R)	Maximum number of exceedance and return period calculations allowed	Pdm.par	Conc	
ERROR	---	Scalar(L)	Used in interactive input	Readr		M
ERRORM	---	Scalar(C)	Error message used in interactive input	Intact		
FILEN	---	Scalar(R)	File name	Openf		
HCONC	---	Scalar(R)	Highest concentration for which return period is desired	Pdm.cmn	Conc	
HDIL	---	Scalar(R)	Calculated dilution at the ALPHA quantile	Moment		
HLINE	---	Scalar(C)	Horizontal line used in interactive input	Intact		
I	---	Scalar(I)	Counter	Intact		
IBEEP	---	Scalar(I)	Flag used in interactive input	Err		I
ICHAR	---	Scalar(I)	Do loop counter	Comrd		
ICONC	---	Array (R)	Increment of concentration for which return period is desired	Pdm.cmn	Conc	
IERR	---	Scalar(I)	Flag indicating error in reading a number	Intact, Readr		O
IFLAG	---	Scalar(I)	Flag indicating error in opening file	Openf		
IFLAG	---	Scalar(I)	Flag indicating error in reading a number	Intact		
IN	---	Scalar(I)	Input file unit number	Comrd		I
INAME	---	Scalar(I)	Number of parameters to read in	Intact		I
INVNRM	---	Scalar(R)	Function which returns the zscore when given the probability	Moment		
IO	---	Scalar(I)	Unit number for file echo	Comrd		I
IOUT	---	Scalar(I)	Unit number of output file	Trnlog		I
IUNUTI	---	Scalar(I)	Unit number used to read from keyboard	Openf		I

TABLE D-1
LIST OF VARIABLES INCLUDED IN THE PDM ANALYTIC MODEL (continued)

Variable Name	Units	Type*	Description	Subroutine /Function /Include File	Common	Code
IUNITO	---	Scalar(I)	Unit number used for output	Main,Openf, Output, Echo		I
IUNITR	---	Scalar(I)	Unit used for input	Main, Readf		I
IVAL	---	Scalar(I)	Error number from opening file (see compiler documentation for explanation)	Openf		
JEND	---	Scalar(I)	Do loop counter	Readf		
JSTRT	---	Scalar(I)	Do loop index	Readf		
LCONC	---	Array (R)	Lowest concentration for which return period is desired	Pdm.cmn	Conc	
LDIL	---	Scalar(R)	Calculated dilution at the (1-ALPHA) quantile	Moment		
LENGTH	---	Scalar(I)	Length of title	Intact		
LINE	---	Scalar(C)	Dummy character variable used for temporary storage	Openf, Readf, Comrd		O
LINEN	---	Scalar(I)	Flag indicating type of box used in interactive input	Main		
LNMEAN	---	Scalar(R)	Mean of distribution in log space	Trnlog		O
LNMEND	---	Scalar(R)	Ln mean of dilution	Main		
LNSTD	---	Scalar(R)	Standard deviation of distribution in log space	Trnlog		O
LNVARD	---	Scalar(R)	Ln standard deviation of dilution	Main, Moment		I
MCOM	---	Scalar(R)	Downstream concentration in arthmetic space	Moment		
MCR	---	Scalar(R)	Mean runoff concentration	Pdm.cmn	Qrcmn	
MCS	---	Scalar(R)	Mean upstream concentration	Pdm.cmn	Qscmn	
MEAN	---	Scalar(R)	Mean of distribution in arthmetic space	Trnlog		I
MESAGE	---	Scalar(C)	Message used in interactive input	Main, Err		I
MNDIL	---	Scalar(R)	LN Mean of dilution	Moment		
MQR	---	Scalar(R)	Mean runoff	Pdm.cmn	Qrcmn	
MQS	---	Scalar(R)	Mean streamflow	Pdm.cmn	Qscmn	

TABLE D-1
LIST OF VARIABLES INCLUDED IN THE PDM ANALYTIC MODEL (continued)

Variable Name	Units	Type*	Description	Subroutine /Function /Include File	Common	Code
MRIM	years	Scalar(R)	Return period	Main, Moment, Output		O, I
NAMES	---	Array (C)	Variable names	Intact		I
NCONC	---	Scalar(I)	Number of exceedance and return periods to calculate	Pdm.cmn	Conc	
NLINEN	---	Scalar(I)	Line number of message written to status box	Trnlog		
NLINES	---	Scalar(I)	Number of lines in the title	Main		
NPERCT	---	Scalar(I)	Maximum number of exceedance probabilities when using method of moments	Main, Pdm.cmn, Output		I
NST	---	Scalar(I)	Number of storms per year	Pdm.cmn	General	
OPTION	---	Scalar(L)	Flag indicating batch or interactive input	Main		
ORDER	---	Scalar(I)	Order of gaussian quadrature used	Pdm.cmn	General	
OUT	---	Scalar(L)	Flag indicating if input is to be echoed as read	Comrd, Readf		I
P	---	Scalar(R)	Probability	Invnrm		
PROB	---	Scalar(R)	Probability P(x) <= normal probability	Invnrm		
PROBLS	---	Array (R)	Probabilities at which to calculate concentration	Pdm.cmn	General	
PROBMR	---	Scalar(R)	1 - PROBLS	Main, Moment, Output		O, I
QUEST	---	Scalar(C)	Character variable containing a question	Readf,Openf, Echo		I
RHO	---	Scalar(R)	Correlation between streamflow and runoff	Pdm.cmn	General	
ROW	---	Scalar(I)	Row on the screen to write a message to	Intact, Openf		
ROWI	---	Scalar(I)	First row on screen used for printing	Intact		
RSTAT	---	Scalar(L)	Logical variable indicating success of read	Main, Readf		O
SDCOM	---	Scalar(R)	Standard deviation of downstream	Moment		O
SDCR	---	Scalar(R)	Standard deviation of runoff concentration concentration in arthmetic space	Pdm.cmn	Qrlcmn	

TABLE D-1
LIST OF VARIABLES INCLUDED IN THE PDM ANALYTIC MODEL (continued)

Variable Name	Units	Type*	Description	Subroutine /Function /Include File	Common	Code
SDCS	---	Scalar(R)	Standard deviation of streamflow concentration	Pdm.cmn	Qslcmn	
SDDIL	---	Scalar(R)	LN standard deviation of dilution	Moment		
SDQR	---	Scalar(R)	Standard deviation of runoff	Pdm.cmn	Qrlcmn	
SDQS	---	Scalar(R)	Standard deviation of streamflow	Pdm.cmn	Qslcmn	
SIGN	---	Scalar(I)	Used to convert answer to proper half of solution domain	Invnrm		
FACTOR	---	Scalar(I)	Used to convert answer to proper half of solution domain	Nrmprb		
STATUS	---	Scalar(C)	Character variable containing statsus of file to open (new or old)	Echo, Readf,	Openf	I
STD	---	Scalar(R)	Standard deviation of distribution in arthemetic space	Trnlog		I
T	---	Scalar(R)	Normalized probability in approximation to gaussian distribution		Invnrm	
TCOM	---	Scalar(R)	Median of downstream concentration in log space	Moment		
TCR	---	Scalar(R)	Median runoff concentration	Pdm.cmn	Qrlcmn	
TCS	---	Scalar(R)	Median streamflow concentration	Pdm.cmn	Qslcmn	
TERM1,TERM2	---	Scalar(R)	Terms used in the solution of inverse gaussian approximation	Invnrm		
TEROFL	---	Scalar(I)	Unit number used for interactive input and	Main, Echo,		I
TITLE	---	Scalar(C)	Character array used to store messages written to screen	Main		
TITLE	---	Scalar(C)	Title of data group	Intact		I
TQR	---	Scalar(R)	Median runoff	Pdm.cmn	Qrlcmn	
TQS	---	Scalar(R)	Median streamflow	Pdm.cmn	Qslcmn	
UCOM	---	Scalar(R)	Mean of downstream concentration in log output	Moment Readf		

TABLE D-1
LIST OF VARIABLES INCLUDED IN THE PDM ANALYTIC MODEL (concluded)

Variable Name	Units	Type*	Description	Subroutine /Function /Include File	Common	Code
UCR	---	Scalar(R)	Ln mean of runoff concentration	Pdm.cmn	Qrlcmn	
UCS	---	Scalar(R)	Ln mean streamflow concentration	Pdm.cmn	Qslcmn	
UD	---	Scalar(R)	Ln mean of dilution	Main		
UQR	---	Scalar(R)	Ln mean of runoff	Pdm.cmn	Qrlcmn	
UQS	---	Scalar(R)	Ln mean streamflow	Pdm.cmn	Qslcmn	
VALUE	---	Array (R)	The value read from the screen.	Readr		M
VALUES	---	Scalar(R)	Values associated with NAMES used in interactive input	Intact		I
WCOM	---	Scalar(R)	Standard deviation of downstream concentration in log space	Moment		
WCR	---	Scalar(R)	Ln standard deviation runoff concentration	Pdm.cmn	Qrlcmn	
WCS	---	Scalar(R)	Ln standard deviation streamflow concentration	Pdm.cmn	Qslcmn	
WD	---	Scalar(R)	Ln standard deviation of dilution	Main		
WIDTH	---	Scalar(I)	Width of screen in columns	Intact		
WQR	---	Scalar(R)	Ln standard deviation runoff	Pdm.cmn	Qrlcmn	
WQS	---	Scalar(R)	Ln standard deviation streamflow	Pdm.cmn	Qslcmn	
ZM	---	Scalar(R)	Zscore corresponding to probabilities in PROBLS	Moment		
ZSCR	---	Scalar(R)	Zscore of normal probability function	Nrmprb		
ZSCR2	---	Scalar(DP)	ZSCR squared	Nrmprb		
ZSCR3	---	Scalar(DP)	ZSCR cubed	Nrmprb		
ZSCR4	---	Scalar(DP)	ZSCR to the fourth power	Nrmprb		
ZSCR5	---	Scalar(DP)	ZSCR to the fifth power	Nrmprb		

* Types are: (I) Integer, (R) Real*4, (DP) Real*8, (C) Character, (L) Logical

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES

Variable	Type*	Description	Subroutine or Function	Designation**
A	Array (DP)	Array to be multiplied by a vector	MTPV	I
A	Scalar (C)	Character string to be left-justified and capitalized	LFTJUS NAMFIX LNGSTR	M M M
A	Scalar (R)	Minimum value for a distribution	TRANSB	I
A	Scalar (R)	Triangular distribution parameter	RANDOM	L
AMN	Scalar (R)	Minimum value of an output data set	OUTFOR FRQTAB	L I
AMX	Array (R)	Maximum value of an output data set	OUTFOR	L
ANS	Scalar (C)	Character string to be capitalized	BMPCHR MAIN	L
ARG	Scalar (R)	Value used to calculate log normal distribution parameters	TRNLOG	L
AV	Scalar (R)	Intermediate value used to generate normal random numbers	NMB	M
B2	Scalar (R)	Triangular distribution parameter	RANDOM	L
B	Array (DP)	Decomposed correlation matrix	RANDOM DECOMP	O M
B	Array (R)	Vector to be multiplied by an array	MTPV	

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
B	Scalar (R)	Maximum value for a distribution	TRANSB	I
BBT	Array (DP)	Correlation matrix for Monte Carlo inputs	READM INITMC STATIS MCECHO DECOMP	O I L I M
C	Scalar (R)	Triangular distribution parameter	RANDOM	L
CN	Array (R)	Array of random correlated numbers	RANDOM TRANSM	M I
CO	Scalar (R)	Downstream concentration	MAIN, PDM	O
CORR	Array (DP)	Array of correlation terms for summary output variables	INITMC STATIS OUTPUT STOUT	M M I M
CORR	Scalar (R)	Correlation coefficient	READM	L
CR	Scalar (R)	Runoff concentration	MAIN, PDM	I
CS	Scalar (R)	Upstream concentration	MAIN, PDM	I
CVAR	Scalar (R)	Coefficient of variation	STOUT	L
CV	Scalar (R)	Coefficient of variation	OUTFOR FRQTAB	L I
DECOM	Array (DP)	Decomposed correlation matrix for Monte Carlo inputs	INITMC	O
DIST	Array (R)	Array storing empirical distributions	READM RANDOM MCECHO EMPCAL	O I I I

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
FACTOR	Scalar (R)	Power of ten values on the X-axis of cumulative distribution plots	OUTFOR FRQPLT	L I
FATAL	Scalar (L)	Variable indicating error in input data	COMRD ERRCHK	L I
FN	Scalar (R)	The number of values	STOUT	L
FX	Scalar (R)	A randomly-generated cumulative probability	EMPCAL RANDOM	L L
FXM	Scalar (R)	Triangular distribution parameter	RANDOM	L
GR	Array (C)	Grid line contents	FRQPLT	L
G	Scalar (R)	Intermediate value used to generate normal random numbers	NMB	M
HH	Scalar (DP)	Element of the decomposed correlation matrix	DECOMP	L
H	Scalar (DP)	Element of the decomposed correlation matrix	DECOMP	L
ICDF	Scalar (I)	Counter for cumulative distributions	OUTPUT OUTFOR	L I
ICHAR	Scalar (I)	Counter for number of characters in a line of data	COMRD	L
ICH	Scalar (I)	The number of characters in a character string	BMPCHR	L
ICOUNT	Scalar (I)	Counter for correlated variables	MCECHO	L

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
IERROR	Scalar (I)	Flag indicating error condition	ERRCHK	I
IERR	Scalar (I)	Input error flag	READM RANDOM MCECHO COMRD TRANSM	L L I L O
IE	Scalar (I)	Counter for empirical distributions	READM MCECHO	L L
IHL	Scalar (C)	"-" character	FRQPLT	L
IJ	Scalar (I)	Do loop counter	FRQPLT	L
IN2	Scalar (I)	Monte Carlo input file unit number	READM	I
INDEX	Scalar (I)	Counter for empirical distributions	RANDOM OUTFOR EMPCAL	L L I
IOUT2	Scalar (I)	Output file unit number for results of each Monte Carlo run	STATIS	I
IOUT	Scalar (I)	Output file unit number for the summary output file	OUTPUT STOUT	I I
IPERC	Scalar (C)	"%" character	FRQPLT	L
IPL	Scalar (C)	"+" character	FRQPLT	L
IRUN	Scalar (I)	Counter for number of Monte Carlo runs	STATIS	I
ISCR	Scalar (I)	Unit number for screen input and output	READM INITMC COMRD ERRCHK DECOMP	I I I I I

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
ISEED	Scalar (I)	Seed for uniform random number generator	RANDOM EXPRN EMPCAL	O M I
ISTAR	Scalar (C)	"*" character	FRQPLT	L
ISTRT	Scalar (I)	Counter for the first non-blank character in a character string	LFTJUS	L
I	Scalar (I)	Do loop counter	READM INITMC RANDOM STATIS OUTPUT MCECHO COMRD LFTJUS NAMFIX MTPV EXPRN DECOMP OUTFOR FRQPLT FRQTAB STOUT EMPCAL	L L L L L L L L L L M L L L L L L
ITYPE	Scalar (I)	Flag indicating whether variables are transformed to or from Johnson SB space	TRANSM TRANSB	L I
IVL	Scalar (C)	"!" character	FRQPLT	L
IWDTH	Scalar (I)	The length of a character string	LFTJUS NAMFIX	L L
IWRTX	Scalar (I)	Format for label on X-axis of cumulative probability plots	OUTFOR FRQPLT	L L
IWRTY	Scalar (I)	Y-axis format index	FRQPLT	L

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
IX	Scalar (I)	The value of the random number used	UNIF NMB FRQPLT	M M L
IY	Scalar (I)	Do loop counter	FRQPLT	L
JE	Scalar (I)	Do loop counter	READM MCECHO	L L
JJ	Scalar (I)	Counter for correlated input variables	READM	L
J	Scalar (I)	Do loop counter	READM INITMC STATIS MCECHO DECOMP OUTFOR FRQPLT FRQTAB STOUT	L L L L L L L L L
K	Scalar (I)	Counter for correlated input variables	READM INITMC MTPV TRANSM DECOMP OUTFOR FRQPLT	L L L L L L L
K1	Scalar (I)	Intermediate integer used in generating uniform random numbers	UNIF	M
KFILE	Scalar (I)	Input file unit number	COMRD	L
KK1	Scalar (I)	Counter used to write names of correlated variables	MCECHO	L
KK2	Scalar (I)	Counter used to write names of correlated variables	MCECHO	L

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
KK	Scalar (I)	Counter for correlated input variables	READM MCECHO	L L
KOUNT	Scalar (I)	Counter for correlated variables	MCECHO	L
L0	Scalar (C)	"0" character	FRQPLT	L
L1	Scalar (C)	"1" character	FRQPLT	L
L2	Scalar (C)	"2" character	FRQPLT	L
L3	Scalar (C)	"3" character	FRQPLT	L
L4	Scalar (C)	"4" character	FRQPLT	L
L5	Scalar (C)	"5" character	FRQPLT	L
L6	Scalar (C)	"6" character	FRQPLT	L
L7	Scalar (C)	"7" character	FRQPLT	L
L8	Scalar (C)	"8" character	FRQPLT	L
LARR	Array (I)	Array storing array addresses for random input variables	READM INITMC MCECHO	M M I
LA	Scalar (C)	"A" character	FRQPLT	L
LC	Scalar (C)	"C" character	FRQPLT	L
LE	Scalar (C)	"E" character	FRQPLT	L
LF	Scalar (C)	"F" character	FRQPLT	L
LINE	Scalar (C)	Temporary storage of the contents of input data lines	READM MCECHO COMRD	L L O
LI	Scalar (C)	"I" character	FRQPLT	L
LL	Scalar (C)	"L" character	FRQPLT	L

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
LM	Scalar (C)	"M" character	FRQPLT	L
LNMEAN	Scalar (R)	The mean of log-transformed data	TRNLOG	O
LNSTD	Scalar (R)	The standard deviation of log-transformed data	TRNLOG	O
LN	Scalar (C)	"N" character	FRQPLT	L
LQ	Scalar (C)	"Q" character	FRQPLT	L
LR	Scalar (C)	"R" character	FRQPLT	L
L	Scalar (I)	Index for an element of the decomposed correlation matrix	DECOMP	L
LT	Scalar (C)	"T" character	FRQPLT	L
LU	Scalar (C)	"U" character	FRQPLT	L
LV	Scalar (C)	"V" character	FRQPLT	L
LY	Scalar (I)	Interval number on Y-axis	FRQPLT	L
LYY	Scalar (C)	"Y" character	FRQPLT	L
MCMAX	Scalar (I)	Maximum possible number of random input variables	READM INITMC MCECHO EMPCAL	I I I I
MCVAR	Scalar (I)	Number of random input variables	READM INITMC MCECHO	O M L
MEAN	Scalar (R)	The arithmetic mean of a variable	TRNLOG OUTFOR FRQTAB	I L I

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
MESAGE	Scalar (C)	Error message to be written to the terminal screen	COMRD ERRCHK	L I
MVAR	Scalar (I)	The number of random input variables	RANDOM TRANSM	I I
MVARX	Scalar (I)	The maximum allowed number of input variables	RANDOM TRANSM	I I
NAME1	Scalar (C)	Label corresponding to the first correlated variable	READM	L
NAME2	Scalar (C)	Label corresponding to the second correlated variable	READM	L
NCARLO	Scalar (I)	The number of random input variables	MTPV DECOMP	I L
NCDF	Scalar (I)	The number of cumulative distribution plots	READM STATIS	L L
NCMAX	Scalar (I)	Maximum number of variables for which cumulative distributions can be plotted	READM STATIS OUTPUT OUTFOR	I I I I
NDAT	Array (I)	Number of values in empirical distributions	READM RANDOM MCECHO	O I I
NDAT	Scalar (I)	Number of values in empirical distributions	EMPCAL	I
NDAGE	Scalar (I)	Number of values in empirical distributions	READM	L

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
NDIS	Array (C)	Probability distribution names	MCECHO	L
NEMP	Scalar (I)	Maximum data pairs allowed to input to empirical distributions	READM RANDOM MCECHO EMPCAL	I I I I
NEVAR	Scalar (I)	The number of empirically-distributed input variables	READM MCECHO	L I
NMAX	Scalar (I)	Maximum possible number of variables for which summary statistics can be printed	READM INITMC STATIS OUTPUT MCECHO STOUT	I I I I I I
NRMAX	Scalar (I)	Maximum number of Monte Carlo simulations allowed	READM STATIS OUTPUT OUTFOR	I I I I
NRUNS	Scalar (I)	The number of Monte Carlo runs	OUTPUT STOUT	I I
NRUN	Scalar (I)	The number of Monte Carlo runs	READM MCECHO	O I
N	Scalar (I)	The dimensions of the decomposed correlation matrix	MTPV DECOMP	I L
NUMPLT	Scalar (I)	Plot number	FRQPLT	L
NVAL	Scalar (I)	The number of values in the data set to be plotted	OUTFOR FRQTAB	I I

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
NVAR	Scalar (I)	Number of summary output variables	READM INITMC STATIS OUTPUT MCECHO STOUT	O I I I I I
NWRITE	Scalar (I)	The number of variables written to an output file for each Monte Carlo run	STATIS	L
NY	Scalar (I)	The number of intervals on the Y-axis	FRQPLT	L
OUT2	Scalar (I)	Output file unit number for results of each Monte Carlo run	READM MCECHO	I I
OUTMSG	Scalar (C)	Error message written to the terminal screen	ERRCHK	L
OUT	Scalar (I)	Monte Carlo summary output file unit number	READM MCECHO	I I
PALPH	Scalar (R)	Probability associated with quantile confidence bounds	READM OUTPUT MCECHO OUTFOR FRQTAB	O I I I I
PC	Array (R)	Percent of time in intervals	OUTFOR FRQPLT FRQTAB	L I I
PCC	Array (R)	Exceedance probabilities	OUTFOR FRQPLT FRQTAB	L I I
PCCM1	Array (R)	Cumulative probabilities	OUTFOR	L
PCZERO	Scalar (R)	Cumulative probability for the first interval	OUTFOR FRQTAB	L I

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
PERC50	Array (R)	The 50th percentile and confidence bounds	OUTFOR FRQTAB	L I
PERC80	Array (R)	The 80th percentile and confidence bounds	OUTFOR FRQTAB	L I
PERC90	Array (R)	The 90th percentile and confidence bounds	OUTFOR FRQTAB	L I
PERC95	Array (R)	The 95th percentile and confidence bounds	OUTFOR FRQTAB	L I
PLTOPT	Scalar (I)	Option flag for cumulative distribution plots	OUTPUT OUTFOR FRQPLT	L I I
PLTTYP	Scalar (I)	Flag for type of plot	FRQPLT	L
PNAME	Array (C)	Input labels used to flag ramdom input variables	READM INITMC MCECHO	O M I
POWER	Scalar (R)	Power of 10 for X-axis	OUTFOR	L
PP	Scalar (DP)	Intermediate value used to decompose the correlation matrix	DECOMP	L
P	Scalar (DP)	Intermediate value used to decompose the correlation matrix	DECOMP	L
PXMAX	Scalar (R)	Maximum value of data set	OUTFOR	L
PYMAX	Scalar (R)	Maximum probability to be plotted	FRQPLT	L
QR	Scalar (R)	Runoff flow	MAIN, PDM	I
QS	Scalar (R)	Upstream concentration	MAIN, PDM	I
R	Array (R)	Array of random values generated from various distributions	RANDOM TRANSM	O O

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
RCV	Scalar (R)	Coefficient of variation	MCECHO	L
RMEAN	Scalar (R)	Mean value of data set	OUTFOR	L
RMN	Scalar (R)	Mean value	MCECHO	L
RVAL	Array (R)	Array containing sorted results of Monte Carlo runs	OUTFOR	I
SD	Scalar (R)	Standard deviation of data set	OUTFOR FRQTAB	L I
SLOPE	Scalar (R)	Slope used to interpolate values from empirical distributions	EMPCAL	L
SNAME	Array (C)	Input labels used to flag summary output variables	READM STATIS OUTPUT MCECHO STOUT	O I I I I
STAT	Array (DP)	Array of summary statistics for output variables	INITMC STATIS OUTPUT STOUT	M M I M
STATUS	Scalar (L)	Variable indicating that the end of the input file has been encountered	READM COMRD	L O
STD	Scalar (R)	The standard deviation of a variable	TRNLOG STOUT	I L
SUM2	Scalar (R)	Summation term	OUTFOR	L
SUM	Scalar (R)	Summation term	OUTFOR	L
T	Scalar (R)	Probability associated with two-tailed confidence bounds	OUTFOR	L

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
TABOPT	Scalar (I)	Option flag for cumulative distribution tables	OUTPUT OUTFOR	L I
TEMP	Scalar (R)	Temporary variable used in sorting data	OUTFOR	L
TEROFL	Scalar (C)	Unit number for plot output	OUTFOR FRQPLT FRQTAB	I I I
TITLE	Scalar (C)	The title of the run	READM MCECHO	L I
TITLE	Scalar (C)	Title for cumulative distribution plots	OUTPUT OUTFOR FRQPLT	L I I
U1	Scalar (R)	Intermediate value used to generate normal random numbers	NMB	M
U2	Scalar (R)	Intermediate value used to generate normal random numbers	NMB	M
U	Scalar (R)	Intermediate value used to generate normal random numbers	NMB	M
VALUE	Scalar (R)	The value interpolated from empirical distributions	EMPCAL	O
VAR	Array (R)	Array storing distribution parameters for random input variables	READM INITMC RANDOM MCECHO TRANSM	O M I I I
V	Scalar (R)	Intermediate value used to generate normal random numbers	NMB	M

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Continued)

Variable	Type*	Description	Subroutine or Function	Designation**
W	Scalar (R)	Width of triangular distribution base	RANDOM NMB	L M
X	Array (R)	Array storing values of summary output variables	STATIS	I
X	Scalar (R)	Value in Johnson SB space	TRANSB	O
X	Scalar (R)	Value of cumulative probability	EXPRN	M
XCDF	Array (R)	Array storing values for plotting cumulative distributions	STATIS OUTPUT	M I
XLAB	Array (R)	Label for X-axis of plot	OUTFOR	L
XLAB	Array (R)	Values for X-axis	FRQPLT	I
XLIM	Array (R)	Values corresponding to exceedance probabilities	OUTFOR FRQTAB	L I
XMAX	Scalar (R)	Maximum value	OUTFOR FRQTAB	L I
XMIN	Scalar (R)	Minimum value	OUTFOR FRQTAB	L I
XM	Scalar (R)	Triangular distribution parameter	RANDOM	L
XNAME	Scalar (C)	Label for the X-axis of the plot	OUTFOR FRQPLT	I I
XNVAL	Scalar (R)	Number of values in data set	OUTFOR	L

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Concluded)

Variable	Type*	Description	Subroutine or Function	Designation**
YMAX	Scalar (R)	Maximum probability to be plotted	OUTFOR FRQPLT	L I
Y	Scalar (R)	A value transformed from Johnson SB space to normal space	TRANSB	O
Z	Scalar (R)	Standard normal variate corresponding to percentile confidence level	OUTFOR	L
<u>Array Dimensions</u>				
MCMAX	Scalar (I)	Maximum possible number of random input variables	Main	
NMAX	Scalar (I)	Maximum possible number of variables for which summary statistics can be printed	Main	
NCMAX	Scalar (I)	Maximum number of variables for which cumulative distributions can be plotted	Main	
NRMAX	Scalar (I)	Maximum number of Monte Carlo simulations allowed	Main	
NEMP	Scalar (I)	Maximum number of data pairs (empirical distribution, probability value) allowed to input the empirical distribution	Main READM, Random	
<u>Monte Carlo Inputs</u>				
BBT	Array (DP)	Correlation matrix for inputs	Main READM, INITMC	
DIST	Array (DP)	Array storing empirical distributions	Main READM, Random	
IN2	Scalar (I)	Monte Carlo input file	Main	

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Concluded)

Variable	Type*	Description	Subroutine or Function	Designation**
IN2	Scalar (I)	Monte Carlo input file unit number	Main READM	
IOUT	Scalar (I)	Monte Carlo summary output file unit number	Main READM, OUTPUT	
IOUT2	Scalar (I)	Output file unit number for results of each Monte Carlo run	Main STATIS	
ISCR	Scalar (I)	Unit number for screen input output	Main READM, INITMC	
MCVAR	Scalar (I)	Number of random input variables	Main READM, INITMC RANDOM	
NDAT	Array (I)	Number of values in empirical distributions	Main READM, RANDOM	
NRUNS	Scalar (I)	Number of Monte Carlo runs	Main READM, OUTPUT	
NVAR	Scalar (I)	Number of summary output variables	Main	
PALPH	Scalar (I)	Probability associated with quantile confidence bounds	Main READM, OUTPUT	
PNAME	Array (C)	Input labels used to flag random input variables	Main READM, INITMC	
SNAME	Array (C)	Input labels used to flag summary output variables	Main READM, OUTPUT	
VAR	Array (R)	Array storing distribution parameters for random input variables	Main READM, INITMC, RANDOM	

Table D-2.
DESCRIPTION, TYPE, AND DESIGNATION OF VARIABLES USED BY MONTE CARLO ROUTINES (Concluded)

Variable	Type*	Description	Subroutine or Function	Designation**
Storage Arrays				
CN	Array (R)	Temporary array of correlated random numbers	Main RANDOM	
CORR	Array (DP)	Array of correlation terms for summary output variables	Main STATIS, OUTPUT	
DECOM	Array (I)	Decomposed correlation matrix for Monte Carlo inputs	Main INITMC, RANDOM	
IRUN	Scalar (I)	Do loop counter for Monte Carlo runs	Main STATIS	
IVAR	Scalar (I)	Do loop counter for variable number	Main	
LARR	Scalar (I)	Array storing array addresses for random input variables	Main READM, INITMC	
RMC	Array (R)	Array of randomly-generated numbers	Main RANDOM	
STAT	Array (DP)	Array of summary statistics for output variables	Main STATIS, OUTPUT	
XCDF	Array (R)	Array storing values of selected variables for plotting cumulative distributions	Main STATIS, OUTPUT	
XMC	Array (CR)	Array storing values of summary output values	Main STATIS	

*DP = Double Precision
I = Integer
R = Real
C = Character
L = Logical

**I,O,M,L, = Input, Output, Modified, Local

☆U.S. Government Printing Office: 1989-625-505