

A STATISTICALLY BASED PROCEDURE FOR CALIBRATION
OF WATER DISTRIBUTION SYSTEMS

By

JOSÉ VICENTE GRANATO DE ARAÚJO

Bachelor of Science
Universidade Federal de Goiás
Goiânia, Brazil
1981

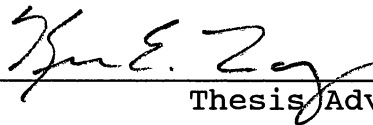
Master of Science
Oklahoma State University
Stillwater, Oklahoma
1989

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1992

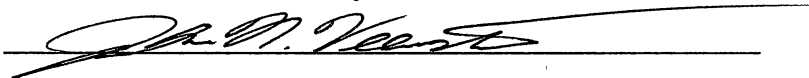
Thesis
1999.0
A6t3S

A STATISTICALLY BASED PROCEDURE FOR CALIBRATION
OF WATER DISTRIBUTION SYSTEMS

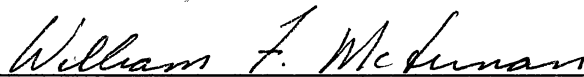
Thesis Approved:



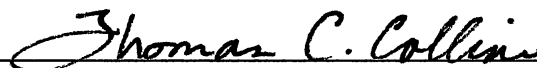
Thesis Adviser











Dean of the Graduate College

ACKNOWLEDGMENT

I would like to express my indebtedness to Dr. Kevin E. Lansey, my thesis advisor, for the inspiration of this research, intelligent guidance and invaluable help. He made me learn and overcome the difficulties of opening a new area of research. His continuous effort, dedication, and more importantly, his friendship made it possible and easier to complete this endeavor. My appreciation is extended to Dr. Charles T. Haan, who most gratefully adopted me in the middle and final stages of the research; for his time, professional experience, guidance, interest, careful readings and critiques to the dissertation.

Thanks to Dr. John Veenstra, chairman of my advisory committee, for all his assistance during the course of my studies at Oklahoma State University. Thanks also to Dr. William F. McTernan and to Dr. Marvin P. Terrell for serving as committee members, for their critiques and for bringing different backgrounds which enriched the development of this work. Thanks to the Civil Engineering Department and to the Water Research Center for their partial support during the development of this research.

Special thanks to my lovely wife, Gisele, for helping me in the mathematical background and in the preparation of the final form of the thesis. She has been a wonderful

companion; her constant smile, moral support, patience, trust and love have made the time of our graduate studies nicer and pleasant.

The pursuit of a doctoral degree presents many challenges especially for a foreign student with limited time to stay abroad. A special thanks to my mother Isa, for taking care of my affairs in Brazil, for her courage, support, dedication to the family, and countless effort in finding special ways to extend my stay in the U.S. to successfully complete my studies. Nothing can exceed my appreciation to her except her love to me. Thanks also to my brothers Marcelo, Roberto, André, Sérgio and Paulo, and my sisters, Maria Inês and Adriana and to their families for understanding the lack of my presence, and for helping my mother during my absence.

Special thoughts go to my late father and professor Eng. Djalma Araujo, who introduced me to the field of Civil Engineering. He left a real testimony of christianity and professional integrity which serves as example and inspiration to my life. It is to his memory that I dedicate this dissertation.

My father-in-law, Dr. Jerson, is recognized for convincing me toward the decision to come to the U.S. for an advanced degree. Thanks also to my mother-in-law, Balbina, for her attention and uncountable letters which kept us well informed of the happenings in Brazil.

My sincere gratitude to the Federal and local State

Brazilian agencies, CAPES, UFG and SANEAGO, for providing with the legal permission and financial support that made possible my coming and study in the USA.

Finally, I want to thanks all unnamed persons who somehow helped me to obtain this degree. The secretaries, support personnel, my colleagues, and my friends here and in Brazil. They are all in part responsible for my accomplishments and I share with them the honor of my success.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
II. BACKGROUND INFORMATION AND REVIEW OF ESTIMATION TECHNIQUES.....	4
Water Distribution Modeling.....	4
Calibration Procedure.....	6
Hydraulic Equations for Distribution Network.....	7
Loop Equations.....	9
Node Equations.....	13
Network Simulator.....	14
Parameter Estimation Techniques.....	15
Analytical Methods.....	17
Optimization Methods.....	20
Uncertainty Analysis.....	22
Other Applications of Parameter Estimation.....	23
Parameterization and Model Reduction.....	24
Data Collection and Optimal Experiment Design.....	25
III. RESEARCH OBJECTIVE AND SCOPE OF WORK.....	27
Scope of the Work.....	27
Uniqueness of Research.....	29
IV. CALIBRATION PROCEDURE.....	31
Basic Steps of the Calibration Methodology.....	31
Calibration Assessment.....	36
First Order Analysis.....	40
Monte Carlo Method.....	41
Data Collection.....	43
Procedures to Improve Calibration Efforts.....	44
Global System Tests.....	45
Identification of Loading Condition.....	49
Criteria to Select the Measurement Load.....	50

Chapter	Page
Identification of Critical Nodes.....	53
Improving Knowledge of Individual Pipes.....	56
Field Observations and Pseudo-Measurements.....	59
Parameter Estimation.....	60
Linear Regression for Coefficient C...	62
Uncertainty in Calibration Data.....	63
Uncertainty in Estimated Parameters...	66
V. APPLICATION OF PARAMETER ESTIMATION AND ASSESSMENT PROCEDURES.....	69
Network Description.....	70
Demand Patterns.....	72
Estimation of Model Parameters with Uncertainty Data.....	74
Calibration Assessment.....	82
Monte Carlo Method.....	85
First Order Analysis.....	87
VI. EVALUATION OF THE ASSESSMENT CRITERIA.....	90
Network Description.....	91
Uncertainties in Nodal Demands.....	94
Uncertainties in Nodal Pressures.....	98
Assessment of Calibration for a Different Demand.....	100
Impact of Different Assessment Loads.....	104
Calibration with Measurements from Low Stress Conditions.....	106
Different Amount of Available Information..	110
Network Parameterization.....	113
Convergence of the Mean and Standard Deviation of C's.....	115
Approximation of the Distribution of Parameters.....	118
Distribution of Computed Pressures.....	121
Appropriateness of the Assessment Measure..	123
Other Useful Descriptive Criteria.....	125
VII. APPLICATION OF DATA COLLECTION PROCEDURE.....	129
System Wide Tests.....	129
Selection of Potential Measurement Loads.....	133
Collection of New Measurements.....	139
Individual Pipe Tests.....	143
VIII. CONCLUSION AND FUTURE WORK.....	147
Conclusion.....	148
Future Work.....	150

Chapter	Page
REFERENCES.....	152
APPENDICES.....	157
APPENDIX A - ADDITIONAL RESULTS FOR CHAPTER V...	158
APPENDIX B - ADDITIONAL RESULTS FOR CHAPTER VI..	161
APPENDIX C - GENERATION OF MULTIVARIATE DEVIATES.....	167
APPENDIX D - PROGRAM LISTING.....	170
APPENDIX E - STOCHASTIC CONVERGENCE OF THE MEAN AND STANDARD DEVIATION OF ROUGHNESS PARAMETERS.....	215
APPENDIX F - PROBABILITY DISTRIBUTION OF ESTIMATED ROUGHNESS PARAMETERS...	232
APPENDIX G - PROBABILITY DISTRIBUTION OF COMPUTED NODAL PRESSURES.....	241

LIST OF TABLES

Table	Page
5.1. Pipe Characteristics for Network 1.....	73
5.2. Available Information Used to Estimate Parameters for Network 1.....	75
5.3. Estimated Pipe Roughness C for Different Levels of Uncertainties in Q's.....	80
5.4. Correlation Among Parameters for CV of Q Equal 5%.....	83
5.5. Assessment Load for Network 1 with Fire at Node 2.....	84
5.6. Calibration Assessment for Network 1 by Monte Carlo Method.....	87
5.7. Calibration Assessment for Network 1 by First Order Method.....	89
6.1. Pipe Characteristics for Network 2.....	93
6.2. Available Information Used to Estimate Parameters for Network 2.....	95
6.3. Estimates of Pipe Roughness C for Different Levels of Uncertainties in Q's and Std of H's Equal 0.25 Feet.....	98
6.4. Estimates of Pipe Roughness C for Different Levels of Uncertainties in H's and CV of Q's Equal 5%.....	99
6.5. Assessment Load for Network 2.....	101
6.6. Calibration Assessment for Network 2 by Monte Carlo Method.....	102
6.7. Calibration Assessment for Network 2 by First Order Method.....	103
6.8. Comparison of Calibration Assessment Computed by Different Methods.....	104

Table	Page
6.9. Comparison of Uncertainty Measures for Different Assessment Loads.....	106
6.10. Comparison of Calibration Performed with Measurements from Different Levels of Network Stress.....	108
6.11. Effects of the Amount of Available Information in the Assessment Measure.....	111
6.12. Effects of Network Parameterization in the Estimates of Roughness Factor for Pipes.....	114
7.1. Estimates of Pipe Roughness C Using Information from 4 Measurement Loads.....	131
7.2. Assessment Load for Data Collection Example...	132
7.3. Calibration Assessment with Parameters Estimated Using 4 Measurement Loads.....	132
7.4. Potential Measurement Loads for Data Collection.....	134
7.5. Selection of the Measurement Load.....	137
7.6. Adjustment of Nodal Demands for the Selected Measurement Load.....	139
7.7. Additional Measurement Load.....	140
7.8. Estimate of Pipe Roughness C Using Information from 5 Measurement Loads.....	141
7.9. Calibration Assessment with Parameters Estimated Using 5 Measurement Loads.....	142
7.10. Comparison of Calibration Assessment Measures for Different Levels of Parameter Uncertainty.....	145

LIST OF FIGURES

Figure		Page
2.1.	Junction Node with Three Pipes.....	14
4.1.	Basic Steps in the Calibration Procedure.....	31
4.2.	Calibration Process Diagram.....	37
4.3.	Selection of Measurement Load.....	47
4.4.	Sensitivity Curves for Pipes.....	59
4.5.	Parameter Estimation Process Diagram.....	64
5.1.	Geometry and Numbering Scheme-Network 1.....	71
6.1.	Geometry and Numbering Scheme for Network 2....	92
6.2.	Stochastic Convergence of Mean C_1	116
6.3.	Stochastic Convergence of Std C_1	117
6.4.	Stochastic Convergence of Mean C_{10}	117
6.5.	Stochastic Convergence of Std C_{10}	118
6.6.	Probability Distribution of C_1	119
6.7.	Probability Distribution of C_3	119
6.8.	Probability Distribution of C_9	120
6.9.	Probability Distribution of H_1	122
6.10.	Probability Distribution of H_4	122

CHAPTER I

INTRODUCTION

Water distribution systems are formed by pipes, pumps, tanks and valves interconnected at junctions or nodes. The distribution network of a municipality can consist of hundreds of pipes or links with decreasing order of importance according to their diameter and location in the system. The hydraulic performance of the water distribution network is generally analyzed through the use of a mathematical model of the system. Such models are representations of the real network and include functional relationships among the components of the network. A network is considered solved when the pressure and consumer demands at all nodes and the flow in the pipes are known (Shamir and Howard, 1977).

Engineers and system operators rely on the model results and use them to make important and costly decisions. Due to the complexity of existing systems and the need to solve problems in a short period of time, the use of simulation models is becoming imperative. Although such models have been used for several decades, the evolution and extensive accessibility of computer hardware and software components makes it possible to apply modeling processes in

every aspect involved with water distribution networks. Model results are used for several purposes including: the design of a new system; the study of system performance under different network stresses; the design of pipe line extensions; the study of rehabilitation alternatives; the analysis of strategies for system operations under normal and emergency situations; etc. Due to the deterioration and aging of existing water systems which were placed in operation many decades ago, it is expected to be invested in the U.S. hundreds of millions of dollars for rehabilitation and expansion of these works.

One of the most important problems concerning the use of mathematical simulators is determining whether the model is actually capable of representing the physical system under study. Proper calibration of model parameters is not an easy task and may be assessed by different methods.

It is of most importance that the model be successfully calibrated before its results can be reliably used for any purpose. The cost of a conservative design increases with increasing uncertainty in pipe roughness (Lansey et al., 1989), so a better calibration will result in less cost for the design. Unfortunately most calibration procedures have been deterministically based (Walski, 1983b, 1986). Little attention has been devoted to the impact of uncertainties in calibration efforts. Additionally, data collection is nearly universally overlooked because it requires great effort and is a costly process. However, this additional

cost may be insignificant if compared to the consequences of making important decisions based on inaccurate results.

This research addresses the problem of improving the calibration of a water distribution network model. It focused on developing a basic framework for the calibration considering uncertainties in input data and model parameters. The calibration procedure incorporates a stochastic component to characterize the several sources of errors and to analyze their propagation through the modeling steps. The objective is to produce a calibrated model which will better represent the physical system. This is accomplished by assessing the uncertainties in model predictions resulting from the uncertainties in input variables and in calibrated parameters.

The calibration assessment determines if the current knowledge of the parameters produces a model which is adequate for the purposes of its use or if it is necessary to collect more information to reduce parameter and model uncertainties. The methodology can also be used to define operational conditions for future data collection to improve the model predictive ability.

CHAPTER II

BACKGROUND INFORMATION AND REVIEW OF ESTIMATION TECHNIQUES

Water Distribution Modeling

A water distribution network is a system formed by a series of pipes which are connected at nodes. Water is supplied to the network from one or more sources named fixed grade nodes. It can be delivered by gravity or pumped into the network. Tanks may also be installed in the system to act as buffers to improve pumping energy efficiency or to function as contingency storage. The energy and flow supplied by pumps and tanks are usually known with only a small degree of uncertainty.

The hydraulic performance of a water distribution network is generally analyzed through the use of a model of the system. Such models are mathematical representations of the real network and include functional relationships among the components of the network. Models are used for several purposes including the design of new systems; the study of system performance under different network stresses; the design of pipe line extensions; the study of rehabilitation alternatives; the analysis of strategies for system operations under normal and emergency situations; etc.

Two major types of parameters must be calibrated in the modeling process of a water network. The first type consists of the roughness coefficients of the pipes which are physical parameters and considered to be time invariant. The other type of parameters is related to the network operational condition consisting of consumer demands and pressure heads at the nodes. These parameters vary in the time domain according to the network usage. Examples of demand conditions are normal, peak and fire loads.

If the nodal demands and pipe characteristics are known, the flow distribution in the network can be computed by solving a system of nonlinear equations. As a system ages, the pipe characteristics, especially the roughness factor, change at an unknown rate and becomes a source of uncertainty to the modeling process.

In order to model a water distribution system, it is convenient to adopt a simplified network to reduce the complexity of the real system. One simplification is to adopt similar pipes. In this case, pipes of the same material, diameter and age are considered to have the same roughness coefficient. However, this coefficient may vary depending upon the pipe's location and amount of flow through their links.

Another modeling practice is to use pipe network skeletonization. In such systems, single pipes are used to represent a group of small links. This may complicate the interpretation of results since the roughnesses of the

modeled pipes may not have a physical basis. The amount of system discretization is a function of the type of study to be developed and depends on the level of detail required for the analysis.

Before a simulation model can be used it must be calibrated to the particular network of concern (Walski, 1983a, 1987). In the calibration process, unknown parameters are adjusted by comparing model predictions with field observations. Once the model is calibrated, it can be used to predict system responses under selected operational conditions.

Calibration Procedure

The calibration procedure is developed based on measurements collected in the field which provide information on system usage and performance. These measurements consist of pressure heads, flow rates, tank levels, valve settings and concentrated demands which are taken at selected points in the system. Nodal demand data usually must be estimated. These data are called pseudo-measurements. They cannot be directly measured because in most cases consumer demands are distributed throughout pipe links. For modeling purposes the distributed flows are lumped at node locations. Lumped demands must be estimated by the modeler or determined in the estimation process.

A typical field test used to collect measurements consists in taking pressure head measurements at selected

points in the system under estimated demand conditions. In a second type of test, one or more fire hydrants are opened causing a fictitious fire demand. Pressure and flow measurements are taken at these points. Remote meters are placed at other locations to record pressure heads during the test. Additional information is available for tanks and pump stations whose pressure heads and flows are monitored.

Besides performing system-wide tests, it is possible to attempt to isolate a single link and conduct a flow test to estimate the pipe roughness. A good estimate of a particular pipes' roughness can be useful in improving the estimation ability of a model.

Hydraulic Equations for Distribution Network

The hydraulic equations of the flow through pipe networks are based in the physical laws of conservation of mass and energy. The first law expresses mass continuity at each junction node. It requires the algebraic sum of flows going into and out of the node to equal zero.

For each junction node j in the system a continuity relationship can be written as (Boulos and Ormsbee, 1991)

$$\sum_{i=1}^{N_p} \lambda_{ji} q_i - Q_j = 0 \quad (2.1)$$

where

N_p = the number of pipes

- λ_{ji} = 0, if the i^{th} pipe is not connected to junction j
 λ_{ji} = 1 or -1, depending whether the i^{th} pipe flow is towards junction j or away from it
 q_i = the volumetric flow rate at pipe i
 Q_j = external demand at junction j .

As water flows in the pipes it loses energy due to the friction with the pipe walls which appears as the pressure drop along the extension of the pipe. The second law requires the conservation of energy along each loop or path in the network. This law is verified when the energy put into the fluid by pumps inside the loop minus the accumulated pipe energy losses sum to zero.

Each primary loop, which is an independent closed path, generates an energy conservation equation of the form (Mays and Tung, 1992)

$$\sum_{i,j \in I_m} h_{L_{ij}} - \sum_{k \in J_m} H_{\text{pump}_k} = 0 \quad (2.2)$$

where

- $h_{L_{ij}}$ = energy loss in pipe linking nodes i and j
 I_m = set of pipes in loop m
 H_{pump_k} = energy introduced by pump k
 m = loop number
 k = refers to pump
 J_m = set of pumps in loop m .

When there is no pump inside the loop, the sum of the energy losses around the loop equals zero. Energy conservation equations can also be written for paths of pipes between any two fixed grade nodes. If there are N_F fixed grade nodes, N_F-1 independent equations can be written as

$$\Delta E_{FGN} = \sum_{i,j \in I_n} h_{L_{ij}} - \sum_{k \in J_n} H_{pump_k} \quad (2.3)$$

where

ΔE_{FGN} = difference in total energy grade between two fixed grade nodes (FGN)

I_n = set of pipes in the path connecting the two FGN

J_n = set of pumps in the same path.

Equation (2.2) can be considered a special case of equation (2.3) where the difference in total grade ΔE is zero for a path which forms a closed loop.

The governing system equations can be expressed in terms of unknown pipe flow or nodal heads using loop (or path) equations or nodal equations.

Loop Equations

In a water distribution network the following equality must hold

$$N_J + N_L + (N_F - 1) = N_P \quad (2.4)$$

where

- N_p = the number of pipes
 N_j = the number of junction nodes
 N_L = the total number of independent loops
 N_F = number of fixed grade nodes.

The structure of the loop equations requires that each component of the path equations, Eq. (2.3), to be expressed in terms of the flow rate. The total energy loss, h_L , in each pipe is due to two sources and may be expressed as

$$h_L = h_{Lp} + h_{Lm} \quad (2.5)$$

where

- h_{Lp} = $K_p q^n$, is the energy loss along the line
 h_{Lm} = $K_m q^2$, is the localized loss due to special fittings
 K_p = a coefficient of physical pipe characteristics
 K_m = coefficient related to the fitting
 q^n = the pipe flow raised to a power n .

Coefficient K_p is a function of the pipe length, diameter and roughness. The computation of K_p as well as the value of the power n depends on whether the energy loss expression used in the analysis is the Hazen-Williams or Colebrook-White equation.

The energy due to the pump in the line may be expressed as a function of its power as

$$H_{pump} = \frac{Z}{q} \quad (2.6)$$

where

- H_{pump} = energy added by the pump
 Z = 8.814 Pu/S (for English units)
 Pu = useful power of the pump
 S = specific gravity of the liquid.

By expressing the energy losses in terms of the flow rate, equation (2.3) becomes

$$\Delta E = \sum (K_p q^n + K_m q^2) - \sum \frac{Z}{q} \quad (2.7)$$

Equations (2.1) and (2.7) form a set of N_p simultaneous nonlinear algebraic equations with respect to unknown flow rates at each pipe. They are called the loop equations.

Typically, the energy relationship for components of water distribution systems is written in terms of the Hazen-Williams equation. Another equation expressing head loss in pipes is the Darcy-Weisbach equation used in conjunction with the Colebrook-White formula. Both flow models have been reported to be well suited for modeling distribution networks under certain conditions (Usman et al., 1988). The Hazen-Williams equation is not very precise in the laminar and transition flow regimes. The Colebrook-White equation incorporates temperature and flow regime dependent coefficients making its use adequate for modeling dynamic situations under varying flow conditions and seasonal variations. Additional review on both equations can be

found in Tullis (1989) and American Society of Civil Engineers (1975).

The Hazen-Williams equation was adopted in this research because of its wide acceptance in the USA and England and pipe flows in most water distribution systems are in the turbulent regime. To represent the energy loss in a pipe the Hazen-Williams equation can be expressed in terms of the flow rate as

$$q = K_1 C_{HW} A R^{0.63} S_f^{0.54} \quad (2.8)$$

where

- q = pipe flow rate in cubic feet per second (cfs)
- K₁ = 1.318 (for English units)
- C_{HW} = Hazen-Williams roughness coefficient
- A = pipe cross section area in square feet (sq-ft)
- R = hydraulic radius of the pipe in feet (ft)
- S_f = h_L/L, slope of the energy line
- h_L = total energy loss between the pipe extremes (ft)
- L = length of the pipe (ft).

Considering no pumps or special fittings in the pipe, the slope of the energy line, S_f, is due to the pipe losses only. The energy loss along the line can be computed as the difference in pressure between junctions at the pipe ends (H_i - H_j). By expressing the area and hydraulic radius as functions of the pipe diameter, equation (2.8) becomes

$$q = K_1 C_{HW} \pi \left[\frac{D^{2.63}}{4^{1.63}} \right] \left[\frac{(H_i - H_j)}{L} \right]^{0.54} \quad (2.9)$$

where

D = pipe diameter used to express the hydraulic radius.

The above equation can be further simplified to

$$Q = \frac{(H_i - H_j)^{0.54}}{K_2} C_{HW} \quad (2.10)$$

where

$$K_2 = 1.755 K_1 \frac{L^{0.54}}{D^{2.63}}$$

The energy loss in the pipe, h_L , can be expressed in a similar form of Eq. (2.5) where n equal 1.852 and K_p equal C_{HW}/K_2 .

Node Equations

The node (or head) equations use only continuity relations defined by Eq. (2.1) and consider nodal pressures as unknown instead of pipe flow rates. For the node represented in Figure 2.1, which has assumed flow directions defined by the arrows, an equation can be written of the form

$$\left(\frac{H_k - H_j}{K_{kj}} \right)^{0.54} + \left(\frac{H_l - H_j}{K_{lj}} \right)^{0.54} - \left(\frac{H_j - H_m}{K_{jm}} \right)^{0.54} - Q_j = 0 \quad (2.11)$$

where

K is the coefficient defined in Eq. (2.5).

Equation (2.11) can be written for each junction in the network forming a system of nonlinear equations with the

number of unknowns corresponding to the number of nodal heads. Since both the loop and node equations represent nonlinear algebraic relationships, no direct solution is possible. They are solved by applying iterative techniques.

Network Simulator

Several algorithms have been developed to solve the nonlinear equations of the flow through pipe networks and have been incorporated in network simulators. One of the most widely accepted of such programs, KYPIPE, the

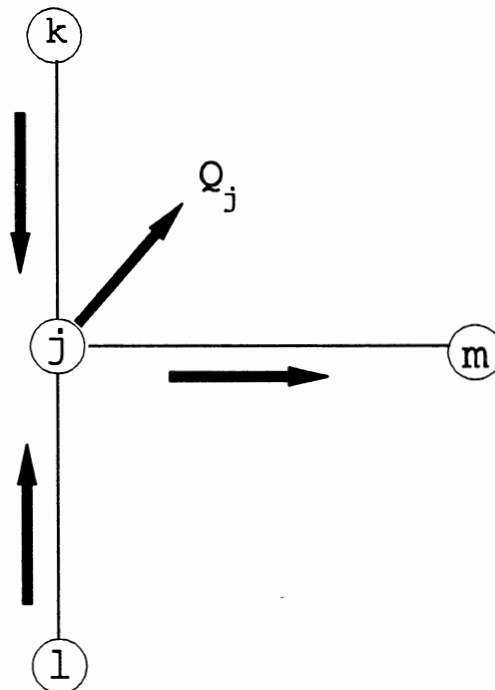


Figure 2.1. Junction Node with Three Pipes

University of Kentucky water distribution simulation model (Wood, 1981), was used in this research. The program uses the linear method that has proved to be very reliable and efficient to solve these equations (Wunderlich and Giles, 1986).

The algorithm used for the solution of the loop equations makes use of gradient methods to handle the nonlinear flow rate (q) terms in the energy equation (2.7). These equations are first linearized in terms of an approximate flow rate q_1 , in each pipe. Based on an arbitrary initial value for the flow in each line, the linearized equations are solved using routine matrix procedures for solving linear equations. A second solution is then obtained. The above procedure is repeated until the changes in flow rates obtained in successive trials are insignificant. Because all flows are computed simultaneously, convergence is expected and occurs within a few iterations.

Parameter Estimation Techniques

Common water distribution system modeling practice is to use ad hoc calibration procedures. Simply, the modeler adjusts pipe roughness coefficients and consumer demands arbitrarily or with judgment to force predictions of pressure heads and pipe flows to agree with measurements taken at a few points in the system.

Walski (1983b, 1986) and Bhave (1988) proposed

equations for correcting C factors and nodal demands based on fire flow tests conducted in the network under both high and low water uses. Rahal et al. (1980) presented a process for adjusting pipe resistance parameters and parabolic pump coefficients for steady state simulations based on the sensitivity of the relative discrepancy between observed and calculated values. The procedure generates global adjustment factors to be applied only to network elements that are affected by the tests. They do not allow calibration of individual elements of the network.

More robust techniques for parameter estimation in water distribution systems may be described as either analytic or optimization methods. Analytic approaches are also referred to as explicit or indirect methods and are frequently used in distribution network systems. Optimization techniques, also referred to as implicit or indirect solutions, have proved to be a powerful tool in the solution of the network equations.

Uncertainty analysis methods, although of frequent use in groundwater and hydrologic systems, have been applied by few researchers as parameter estimation procedures for distribution networks. While the first two approaches are characterized as dealing only with exact values, these methods attempt to analyze the measurement, parameter and model uncertainties. An overview of the parameter estimation methods is presented below.

Analytical Methods

The parameter estimation methods described in this category solve the mathematical formulation of the hydraulics of the water distribution network expressed as a deterministic system of nonlinear equations. This method is commonly used in network simulators considering known values of consumer demands and pipe resistance factors. The system is solved for the unknown pressure at nodes or flow rates in the pipes. The solution methodology consists of inverting the coefficient matrix of the parameters (or using its pseudo inverse). By redefining the unknowns and including field measurements as known information, it is possible to reformulate the equations and solve directly for other unknowns, such as pipe roughness coefficients.

Shamir and Howard (1968) presented a method for solving the network node equations for a combination of unknown elements such as nodal heads, consumptions or pipe resistance factors. A first order Taylor expansion about a current state vector estimate is used to linearize the flow equations which are solved iteratively by the Newton-Raphson method. The system is solved for correction terms of the different unknown elements. The solution of the set of equations is dependent on the distribution of the unknown parameters in the physical system. A unique solution is guaranteed if the matrix formed by the coefficients of the unknown parameters and the constant vector are of full rank.

This criterion can be met if each node in the system has one unknown. The unknowns may be: (1) the external demand at the node, (2) the pressure head at the node or at adjacent node, or (3) the loss coefficient of a pipe which is connected to the node.

Observability analysis for distribution networks attempts to determine a minimum set of measurements that will provide sufficient information for the computation of all other unknown parameters. Bargiela (1985) presented an algorithm for determining observability in water distribution system state estimation (i.e., demand estimation with known pipe parameters), based on the structure of a global matrix defined as a function of meter placement and network topology. His study introduces the concept of topological observability with respect to a given measurement set and suggests a method of finding an observable spanning tree of the network based on graph theory.

Measurements are often scarce in water distribution collection efforts and it is necessary to add pseudo-measurements so that the system of equations remains solvable. The pseudo-measurements introduced are obtained by estimation or interpolation between measurements. These estimates contribute to increase the degree of uncertainty in the calibration process and to reduce model prediction accuracy.

Ormsbee and Wood (1986) developed an explicit method

for the network calibration in which they solve the flow equations in terms of head loss coefficients. The original loop equation formulation considered a system of N_p equations which were solved for N_p unknown pipe flow rates. In the algorithm additional continuity and energy equations supplement the loop equations and allow an equal number of decision variables to be computed. The added variables can be either headloss adjustment factors or head loss calibration coefficients. The solution technique was based on the linear method and was capable of analyzing only a single loading condition.

Cohen and Carpentier (1988) studied observability improvements by taking the dynamic behavior of the system into consideration. They assumed that a static demand condition could be measured at different times with measurements taken at different locations each time. This practice increases the number of measurements for each demand condition. With the additional measurements, the observability criterion may be met. The authors state that the main difficulty of the approach lies in the proper choice of the pseudo-measurements which must be selected according to the particular system.

Parameter estimation under a single loading condition is not a reliable procedure due to compensating errors which might occur in setting parameter values, i.e., the solution is not unique. Several parameter estimation algorithms have been reported to improve calibration of distribution

networks considering multiple demand loads. Rahal and Sterling (1981) considered a dynamic approach for the calibration problem using data from extended period simulations. In this case the loading conditions varied with time and the dynamics of the reservoir hydraulics were included in the solution methodology.

Optimization Methods

Some authors have proposed implicit solutions to estimate parameters of hydraulic network models, based on mathematical programming. Such techniques seek the best values of parameters which minimizes an objective function expressed as a norm of the discrepancy between observed and computed values. These values can be nodal pressures, pipe flow rates, or tank levels, or any combination of the measurements. The procedure also allows one to impose physical limitations on parameters and to fix operating ranges for individual network components which are introduced in a form of constraints in the optimization problem.

Powell et al. (1988) and Sterling and Powell (1989) applied an iteratively re-weighted least squares algorithm for the state estimation problem in which the weights are modified within each Newton-Raphson step to improve convergence. This method, although adequate for on-line monitoring, requires measurement redundancy levels that are not often obtained even in the best telemetry systems. It

also requires previous knowledge of the covariance matrix of the measurement error term which is used as the weight function. This method, however, did not address any uncertainty in estimation of consumer demands or nodal pressure head measurements. Since a linear regression analysis was performed at each step of the analysis, no limit was placed on the acceptable range of the demands.

Ormsbee (1989) and Lansey and Basnet (1991) used nonlinear programming algorithms to estimate demands and pipe roughness coefficients considering individual and multiple loading conditions. Both authors used a problem reduction technique in which the number of constraints were decreased by the use of a simulation model. The simulation model is capable of solving the network equations for the unknown nodal pressure heads, pipe flows and/or tank levels, given a set of roughnesses coefficients and consumer demands which are the parameters to be calibrated. Simulation results are used to evaluate both the objective function and the implicit constraints for a set of calibration parameters generated at each optimization problem iteration.

Lansey and Basnet's procedure allowed one to consider one or more demand pattern and extended period simulation simultaneously. The model's ability of finding optimal solutions is restricted by the type and number of measurements available and by the system observability. Boulus and Ormsbee (1991) extended Ormsbee's work to multiple demands and time varying conditions.

Uncertainty Analysis

The procedures described earlier did not consider the quality of the input values used for the calibration algorithm. They all attempt to find unique values for parameters without assessing their reliability. In a more realistic environment there are several sources of uncertainty present in the various phases of the parameter estimation process.

As discussed earlier the conceptual model of the network is itself a simplification of the real system being modeled. Measurements of heads and flow may contain errors due to the instrument precision and data acquisition and processing. Consumer usage may be over or underestimated. This reflects a disadvantage of the use of a deterministic estimator in that the resulting parameters may be invalid because errors in input data are not accounted for. The use of a purely deterministic estimator may compromise the application of the modeling results due to a lack of knowledge of the reliability of the parameter values and resulting model predictions.

Few authors have addressed uncertainty considerations in modeling water distribution systems. Bargiela(1989) studied the propagation of uncertainty in measurements of pressure heads and estimation of consumer demands to other nodes in the network. He proposed methods for computing error bounds for the demands considering a complete

knowledge of the pipe roughness coefficients. The procedure, however, does not provide any information on the mean, variance and type of distribution for the error term nor does it address the issue of pipe roughness calibration.

Leroy (1988) applied a simple Monte Carlo to examine uncertainty in pipe roughness which consisted basically of performing a sensitivity analysis for the pressure at a few points in the system, considering a highly parameterized network.

Lansey and Basnet (1991) introduced errors in measurements and pseudo-measurements at selected nodes and attempted to calibrate pipe roughness and remaining demands using their gradient based optimization procedure. They did not, however, characterize the structure of the error term or assess the impact of the different magnitudes and types of errors.

Other Applications of Parameter Estimation

Statistical analysis has been applied to study the influence of the errors in the modeling of hydrologic systems (Troutman, 1985). Three types of uncertainties were identified: model errors, input errors, and parameter errors. The error was characterized as a random variable associated with a probability distribution. The identification of the error term was restricted to the knowledge of the distribution of the parameters.

A survey of parameter identification procedures used in groundwater systems is presented by Yeh (1986). The solution methodologies are classified into two broad categories based on the equation error criterion and on the output error criterion. Equation error criteria are the procedures using analytic solutions. In this category model parameters appear as dependent variables and missing data are estimated to make the system solvable. Output error criteria are based on optimization algorithms which minimize a function containing the discrepancies between measured and computed values. A potential problem of not achieving a global optimum may arise since the minimization problem is usually nonlinear and nonconvex. The methods applied in water distribution analysis are similar in sophistication to those presented for modeling groundwater systems.

Parameterization and Model Reduction

Since system parameters must typically be inferred from few measurements, model simplification is commonly adopted. Usual practice is to remove pipes which are assumed to have small influence on system pressures and thus form a smaller network model (skeletonization) or assume that sets of pipes have the same roughness coefficient (parameterization). If pipe characteristics are known, then simple configurations (pipes in series or parallel) can be easily reduced to an equivalent pipe. Hamberg and Shamir (1988) developed a skeletonization procedure which reduces complex sets of

pipes to an equivalent system. The method, however, requires the pipe and demand characteristics to be known.

When the system characteristics are being determined, the above methods cannot be applied without assumptions. Yeh and Yoon (1981) developed a method to examine the impact of parameterization in groundwater systems. Model error which increases with higher levels of parameterization and the parameter error which conversely decreases were computed assuming the complete system is known. The modeler can then define the acceptable tradeoff level.

Data Collection and Optimal Experiment Design

In many systems the structural configuration cannot be directly measured. They are inferred from limited observations of the system responses. These observations must be collected in a time and space domain and they contain errors. A need for an efficient data collection procedure for such systems motivated researchers to consider the problem of experimental design. Optimal experimental design is an active research area in control theory (Walter, 1987). These techniques have been applied and extended in the groundwater area.

The goal of an experimental design is to collect the most valuable information which will reduce the parameter uncertainty and improve the predictive ability of the simulation model. Several optimality criteria have been

applied in designing data collection for groundwater systems. They differ according to the measure of the system reliability adopted. The A-optimality criterion minimizes a trace of the covariance matrix of estimated parameters while satisfying all system constraints. The D-optimality criterion minimizes the determinant of the estimates' covariance matrix. These criteria had been used for scheduling pumping tests for identification purposes (Nishikawa and Yeh, 1989). Other criterion using the concept of maximum information is reported for determining placement of sampling wells for transport parameter identification (Cleveland and Yeh, 1990, 1991).

The objective of these methods is to reduce parameter uncertainty which is not the overriding goal of the modeling effort. Several extended identifiability criteria have recently been introduced which focus on the prediction uncertainty (Chavent, 1987; Hsu and Yeh, 1989; and Sun and Yeh, 1990). These measures have been applied in groundwater estimation to select static experimental designs.

The original work of Federov (1972) introduces the basic concepts of experimental design and presents several identifiability criteria. General algorithms for obtaining D-optimal designs were revised by St. John and Draper (1975).

CHAPTER III

RESEARCH OBJECTIVE AND SCOPE OF WORK

The problem addressed in this research is to improve water distribution system modeling. It is desired to produce a model which best reproduces the responses of the physical system. This requires an understanding of the sources of uncertainties and how they propagate through the modeling process. The overall objective is to develop a framework for calibration of a model to the network system to achieve a desired level of accuracy in the model output values. The goal is to provide engineers and system operators with a reliable system representation allowing them to make more informed decisions.

Scope of the Work

The product of this research is a methodology for calibration of a model for water distribution network systems considering parameter estimation, data collection and techniques for analyzing the parameter estimates' and model predictions' uncertainties. To achieve the research objectives, a series of statistical procedures were linked to consider the effect of measurement uncertainty on the parameter estimates, to assess the impact of parameter

uncertainty on model prediction, and to assist in defining data collection conditions based on model prediction uncertainty. The additional information will be used to reduce parameter and model uncertainties.

The parameters to be estimated are the pipe roughness coefficients for a single pipe or a group of pipes in a skeletonized network. It was assumed that these parameters represent the total loss in each component including minor valve losses and reductions in pipe diameter.

The collected field data are the nodal pressure heads, tank levels, and pump pressure and discharge. Nodal demands are assumed to be known exactly or their statistical distributions are known. Since data collection is expensive and time-consuming, ineffective measurements must be avoided.

Specifically, the research consists of formulating a calibration procedure which considers errors in field measurements and their resulting impact on model predictions. To complete the calibration methodology the components developed were:

1. A linear regression model for estimating system parameters using assumed and/or measured field data and their statistical distributions.
2. An error analysis procedure for the regressed parameters using conditioned Monte Carlo analysis to determine their statistics and distributions.
3. A technique for evaluating the propagation of parameter

errors to model predictive ability using Monte Carlo and first order second moment approaches.

4. A procedure to analyze the worth of future data collection conditions.

The resulting computer programs were then applied to various systems and for different system conditions to learn about the relative importance of the different information in the overall calibration process.

Uniqueness of Research

Most water distribution system models have been deterministically based and do not consider uncertainties introduced by modeling assumptions or errors in parameters. Also, methods for statistically evaluating the goodness of the modeling results are limited.

Current calibration techniques for water distribution networks are insufficient to verify that the model represents reality. Representing reality is the basis upon which all models must be judged and tools to assist modelers achieve these goals are lacking.

This research is an initial effort to quantify the uncertainty in parameter estimates and model predictions. The integrated calibration approach provides guidance in collecting additional data to reduce these uncertainties. The result of applying this methodology will be system models which more closely represent reality. Better models

will lead to more informed decisions regarding operations, design, maintenance and monitoring.

This research is unique for the following reasons.

1. No other published work has considered a comprehensive integrated process for calibrating a water distribution network.
2. This effort is the first to consider the impacts of errors in field data on the model parameters and model predictive ability.
3. No other modeler has considered the parameters and model predictions of water distribution systems as statistical variables which can be represented by their mean and variance, and approximated by a probability distribution.
4. This is the only procedure which examines the value of field information and uses that information to determine useful conditions to guide future data collection efforts.

CHAPTER IV

CALIBRATION PROCEDURE

Basic Steps of the Calibration Methodology

The methodology for the calibration of water distribution systems considering uncertainties is comprised of three basic steps: Data Collection, Parameter Estimation, and Assessment of the calibration. These steps are depicted in Figure 4.1.

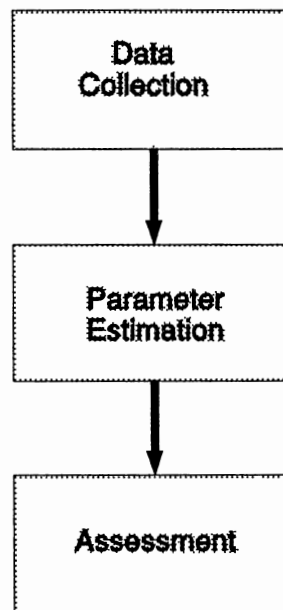


Figure 4.1. Basic Steps in the Calibration Procedure

The purpose of a data collection effort is to provide sufficient and reliable field measurements to be used as input data in the estimation of model parameters. Several types of measurements can be collected at selected locations in the system. These measurements can be the pressure head at junctions, pipe flow rates, and a set of operating conditions on pump status and tank levels reflecting the system usage at the time the measurements were taken. So, additional information on estimation of consumer demands, pump discharges, and valve settings are needed to augment the real measurements and to complement the input data required to estimate the unknown parameters.

Before being accepted as a valid information and used to estimate model parameters, the collected data needs to be filtered to identify and reject gross errors. This screening process determines if the measurements satisfy the physical constraints of the system. If too much information is rejected, observability problems may occur, since it may not be possible to collect the missing data again because the system's operating condition will likely have changed. The observability problem appears when the remaining information is not sufficient for stability of the estimation routine.

The second step in the calibration procedure consists of estimating the unknown parameters. This process aims to determine the best values for the parameters such that model outputs agree with field measurements. In water

distribution systems a typical parameter is the roughness coefficient of pipes, specifically the C-factor in the Hazen-Williams equation. This parameter describes the loss of energy caused by the friction between flowing water with the pipe walls.

Another parameter that must be calibrated is consumer demands which are lumped at node locations. The magnitude of these demands follow a cyclic pattern over the day and also vary with season, weather condition, and emergency situation. Examples of such situations are: high demands used to fight fires, pipe line breaks and pump shutoff caused by power failure. Although a parameter estimation procedure for a water distribution system may allow one to consider other types of unknowns, the procedure developed in this research concentrated primarily on determining values for unknown roughness coefficients. The demands and their statistics were generally assumed known values.

Calibration alone does not provide a means of evaluating how well the model results reproduce the actual system behavior. Evaluation is usually accomplished by validating model results with field measurements which were not used for parameter estimation. To supply information for future data collection efforts an alternative is used in this research.

The third step of the calibration procedure consists of assessing the impact of the calibrated parameters on model predictions. Assessment analysis will determine the

uncertainty in predicting system performance for critical demand conditions. The assessment is usually performed for a different loading condition than those used for calibration. A measure of the calibration assessment reliability must be defined for the specific system and purpose of the simulation model. For example, the variance of the computed pressure at nodes, $\text{Var}(H_a)$, can be used to quantify the uncertainties in model outputs for individual junctions. An overall system measure can be defined as the summation of individual pressure variance at selected nodes. A vector of weights could also be adopted to stress the importance of critical nodes. Another assessment measure could be a norm of the covariance matrix of computed nodal pressures. If the trace of that matrix is adopted as a norm ($\text{trace cov}(H_a)$), the measure would be computed by adding the pressure variances at all nodes.

The reliability measure can be used as the criterion in a procedure to evaluate and select data collection strategies for taking new measurements. An algorithm can be formulated within the calibration loop to minimize this measure for all or selected nodes in the system. The objective of this procedure is to identify conditions that provide the maximum information to obtain reliable estimates of model parameters. It may be found that parameters are not being estimated with the required accuracy to assure a desired level of confidence in the model results. If this happens, the entire process is repeated from step one in

which more data is collected, possibly, under different conditions. The new measurements are expected to augment the available input data in such a way to improve parameter estimates.

The use of more reliable parameters contribute to decrease the uncertainty levels of the results derived from the model. A straight forward validation is also possible to compare model results with field observed conditions. As for any modeling process, validation and calibration should continue to be performed over time after the initial calibration effort is completed.

The assessment of the accuracy of the calibration results is important for design purposes especially in the analysis of nodes with low pressure. A larger variability in the pressure at these nodes is of great concern since there is a risk of not meeting the minimum required pressure for these nodes. When using simulation results to help in operation decisions, it would appear advantageous to have similar pressure uncertainties for all nodes and to have small uncertainty in tank level predictions. In this research the assessment was limited to steady state conditions but it can be extended to consider extended analysis through time.

The methodology for calibration of water distribution systems implements the basic steps of the modeling process discussed: data collection, parameter estimation and calibration assessment. A calibration process diagram

showing the steps involved in the methodology is presented in Figure 4.2. Each block in the diagram represents a specific task which may be accomplished by a single routine or by a sequence of procedures. There may be alternative options to accomplish the tasks specified within each block, as described in detail in the next sections.

As a preliminary step, initial estimates of parameter C's and their variances are assumed based on the available information of the physical system. At this stage a large variance may be assigned to the C's reflecting the high level of uncertainty in their estimates. This step corresponds to block 0 in the calibration diagram. Block 0 is only executed at the initial stage of the calibration effort. After one iteration is completed, parameter estimates are obtained from blocks 10 and 11.

The purpose of a water distribution model is to represent the system so knowledgeable decisions can be made. Therefore, after the parameters and their uncertainties have been estimated, the impact of model reliability must be assessed.

Calibration Assessment

The calibration procedure starts in block 1 by selecting the network loading condition to be reproduced by the model when its calibration has been successfully accomplished. The methodology provides a means to evaluate the calibration accuracy by assessing the model's capability

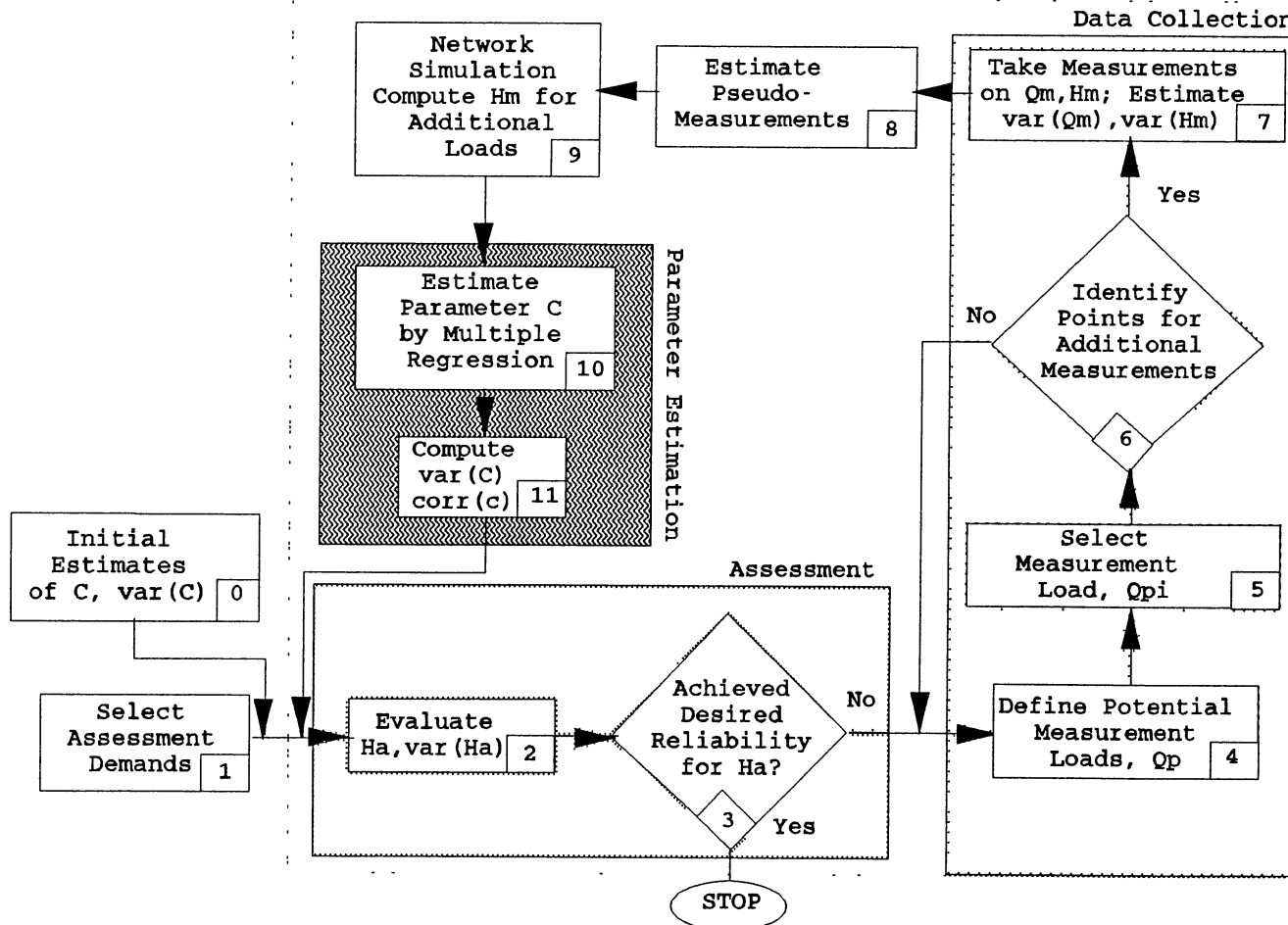


Figure 4.2. Calibration Process Diagram

of simulating the desired network condition (or conditions) such that the uncertainty in model outputs fall within a prescribed range. This range is determined based on the specific purpose of the application under consideration. The accepted range will define stopping criteria for the calibration loop. The selected loading condition, or set of conditions, are called the assessment demands. Usually, assessment demands contain high flows to stress critical nodes in the system. These demands are selected by assigning consumer withdrawals for a single period of time or for a sequence of patterns representing a typical period or an entire day.

In the following module, corresponding to blocks 2 and 3, the calibration assessment is performed for the network conditions specified in block 1 using the current knowledge of the parameters and their level of uncertainty. The assessment consists of looking at prediction uncertainties (or model outputs), in this case, the pressure heads at the nodes. Uncertainties are quantified by an adopted measure which will vary according to the specific calibration objectives and can be either related to the entire system or to individual nodes in the network. This measure can be used in the assessment module as a stopping criteria for the calibration loop. Examples of assessment measures for different calibration objectives are:

1. Single system variance: Trace of the covariance matrix of the simulated nodal pressures for a single assessment

- load;
2. Composite system variance: Composite trace of the covariance matrix of the simulated nodal pressures considering more than one assessment load.
 3. Individual node variance: Variance of the simulated pressure head at a selected node for a single assessment load.
 4. Composite node variance: Composite variance of the simulated pressure heads at a selected node for a series of independent assessment loads.
 5. Multiple node variance: Weighted sum of the variances of the simulated pressures at selected nodes for a single assessment load.
 6. Composite multiple node variance: Composite pressure variance for several nodes for more than one assessment load.
 7. Maximum node variance: Maximum pressure variance at a node or a group of nodes during an extended period of time.

To examine a model's ability to predict nodal pressures, first order second moment analysis or Monte Carlo analysis can be used for estimating variances. The distinctions between these procedures are computational time and accuracy. First order estimates require computing the gradients of the model output with respect to the parameters. A numerical gradient evaluation was applied to

compute the gradients in this research. A Monte Carlo analysis is time intensive particularly for complex network systems under a high degree of uncertainty. Its advantage is that it allows one to obtain some understanding of the distribution of the output which may be used to place confidence intervals on the model results.

First Order Analysis

The first order second moment approach (FOSM) estimates the mean, variance and covariance of model output by approximating the function with a Taylor series expansion around the mean value of the parameters and dropping the higher order terms (Benjamin and Cornell, 1970).

Considering only uncertainties due to the C parameter, the model output uncertainties are computed as

$$\underline{cov(Ha)} = \frac{\partial Ha}{\partial C} \underline{cov(C)} \left[\frac{\partial Ha}{\partial C} \right]^T \quad (4.1)$$

where

$\underline{cov(Ha)}$ = $N_j \times N_j$ covariance matrix of model outputs

$\frac{\partial Ha}{\partial C}$ = $N_j \times N_p$ matrix of partial derivatives of \underline{Ha}
evaluated at the mean value of the C's

$\underline{cov(C)}$ = $N_p \times N_p$ covariance matrix of parameters

\underline{Ha} = $N_j \times 1$ vector of model outputs

\underline{C} = $N_p \times 1$ vector of model parameters

N_j = number of junction nodes in the network

N_p = number of pipes in the network

$[\quad]^T$ = transpose of matrix inside parenthesis.

Monte Carlo Method

The Monte Carlo approach consists of performing network hydraulic simulations repeatedly to compute the mean nodal pressures H_a , and the $\text{cov}(H_a)$ matrix. At each simulation a vector of noise (measurement errors) is added to the mean value of the C parameters to account for the uncertainties in their estimates. The variability in the model output results from the uncertainties in the C parameters. When the procedure is executed for the first iteration, variances of C are assigned based on experience and knowledge of the system.

Parameter correlation can result from field conditions (real correlation) or from the estimation procedure (statistical correlation). Both assessment methods have the capability to reproduce the parameter correlation. In the FOSM approach the $\text{cov}(C)$ is used instead of $\text{var}(C)$ to account for parameter correlation (equation 4.1). In the Monte Carlo method, multivariate normal deviates are generated to introduce errors in the C's which preserve the mean, standard deviation and correlation of parameter C's. The procedure to generate multivariate deviates is shown in Appendix C. It is similar to the procedure used by Borah and Haan (1991) to generate correlated random errors in a hydrologic modeling.

The above methods have different input requirements and

assumptions. The first order, second moment approach requires an assumption on the distribution of model outputs. This assumption become less valid as system nonlinearity increases. It does not require an initial assumption on the distribution of the parameters and it requires few computations. The Monte Carlo method requires fewer assumptions on the transferring of uncertainties through the model. However, it requires the knowledge of the multivariate distribution of the input parameters. Also, it usually requires large computational effort. The purpose of using both methods is to compare the efficiency versus accuracy of the FOSM method of approximating the $\text{var}(H_a)$.

After computing the variance of H_a , the modeler must decide if the calibrated model is satisfactory for the intended use of the simulation results. This condition can be verified by adopting a measure to quantify model output uncertainties. One such measure can be the trace of the covariance matrix of model outputs, in this case, the computed pressures at the nodes. Other prediction error measures could be used such as the ones previously listed in the Parameter Estimation section. Assessment prediction errors based on the standard error, mean error, and mean absolute error have been used by Yeh (1987) to compare different approaches to estimate aquifer parameters. Yan and Haan (1991) used the trace and the determinant of the covariance matrix of the errors as a criterion to estimate parameters for a multiobjective hydrologic model.

If the reliability level is acceptable, the process stops and the estimates of the C's are the values obtained from blocks 10 and 11 in the last iteration. If the measure is not accepted in block 3, it is necessary to decrease the uncertainty in model outputs by improving the knowledge of the estimated parameters. This requires the use of additional, and/or more reliable, field measurements as available information for the estimation procedure.

Data Collection

Before new data can be collected the modeler must define the demand conditions which will provide the maximum amount of information to improve the system's reliability measure. This step is developed in the Data Collection module corresponding to blocks 4 through 8 in the diagram depicted in Figure 4.2.

The objective of block 4 is to define potential demands which should be induced in the network when collecting the new measurements. These demands cannot be arbitrarily selected. They must be achievable in the real system during the period of the data collection experiments. A sequence of demands and tank levels can also be used as additional information.

Upon defining potential measurement loads, the model must be executed to compute the predicted nodal pressures and their reliabilities for each potential load. It is desired to identify the load which will provide the most

valuable information to achieve the calibration objectives. This task is accomplished in block 5 and will be explained in detail later in this chapter. The following step, represented by block 6, will identify potential locations in the network where pressure head measurements should be taken. This is done by selecting the nodes with larger pressure variability. There may be occasions where no point in the system can be measured with a higher degree of accuracy than the one already obtained from block 5. If this is the case, the predicted load is not valuable in providing additional information to improve parameter estimates and another load must be tried. This causes the procedure to return to block 4. A detailed explanation of a process of improving calibration efforts is presented below.

Procedures to Improve Calibration

Efforts

Since data collection is a costly and time consuming process, it must be conducted in the most efficient way. The objective of the data collection module is to provide guidance for obtaining additional information to be used in the calibration algorithm. At each iteration of the calibration process, new measurements are added to the existing data set and used to estimate the unknown parameters.

Alternatively, the new measurements may replace existing and less accurate data so that the available

information will assure the unknown parameters are estimated with less uncertainty. It is expected that the use of parameters with less variability will improve the quality of the simulation. These improvements can be quantified by assessing the uncertainties in the model results, the nodal pressure heads.

Two procedures are presented to improve the calibration efforts. The first procedure considers all system components as potentially active for improving the calibration objectives. The second procedure attempts to enhance the calibration by improving the knowledge of individual network elements such as the pipe roughnesses.

Global System Tests

To improve existing knowledge of all model parameters and consequently to improve the quality of the simulation results, system-wide tests can be conducted to obtain new information. The problem is to determine the best (or at least a worthwhile) loading condition (how much flow to induce and at which node) and to identify measurement type and locations that will be the most beneficial for providing information for the parameter estimation algorithm. This study focuses on identifying the demand conditions.

Sensitivity vectors can be developed to identify potential network conditions and locations to stress the system demands. A fire test is a common way of stressing the distribution network in such a way to mimic an emergency

condition that affects the available flow and pressure at many nodes in the system. A measure related to the calibration objective must be adopted to guide the selection of the location, or locations, which would be the most beneficial for the additional field tests. One such measure could be the total system variance with respect to the simulated pressures. This measure is the trace of the covariance matrix of the simulated nodal pressures and is computed by taking the sum of the variances of the simulated pressures at the nodes.

The procedure to determine field measurement conditions consists of two parts. The first is to select a network loading pattern under which measurements are taken. The second step is to identify the critical nodes where demands must be accurately determined. This can be done by installing flow meters if the demand is localized or by conducting a detailed estimation of the distributed demand which is lumped at the node. The complete procedure is illustrated in Figure 4.3 and will be discussed in the following sections. The sensitivity vectors to be used to identify the ideal loading condition are:

(a) Sensitivity Vector A

$$[\Delta \text{trace cov}(H_p)_1, \dots, \Delta \text{trace cov}(H_p)_i, \dots, \Delta \text{trace cov}(H_p)_L]$$

where

$$\Delta \text{trace cov}(H_p)_i = \text{trace cov}(H_2) - \text{trace cov}(H_p)_i$$

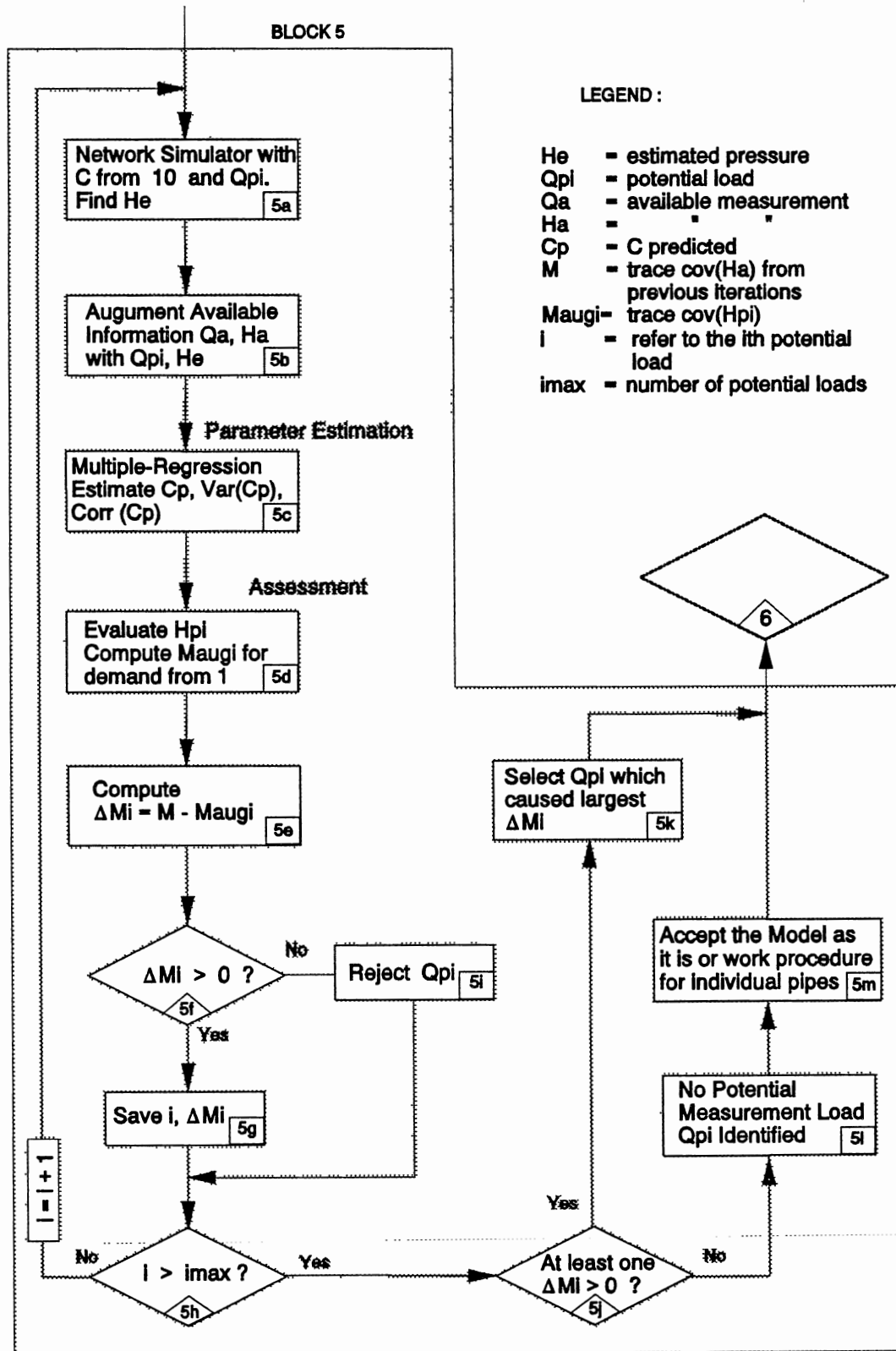


Figure 4.3. Selection of Measurement Load

Each term $\text{trace cov}(H_p)_i$ corresponds to a prediction of the assessment measure to be obtained if the potential measurement load Q_{pi} would be incorporated with the available information to estimate parameters. The trace of $\text{cov}(H_a)$ is the assessment measure at the current stage of the calibration. The modeler wants to improve this measure by collecting new field data under one of the Q_{pi} conditions. L is the total number of such conditions. The potential loading condition Q_{pi} represents a vector containing individual nodal demands.

(b) Sensitivity Vector B

After selecting the most sensitive loading condition, the critical nodes can be identified by examining sensitivity vector B. For a selected node i , this vector is expressed by

$$[\Delta \text{trace cov}(H_p)_{i1}, \dots, \Delta \text{trace cov}(H_p)_{ij}, \dots, \Delta \text{trace cov}(H_p)_{iN_j}]$$

Each term of the sensitivity vector B corresponds to the difference of the trace of $\text{cov}(H_a)$ and the trace obtained if the potential measurement load Q_{pi} is used to estimate parameters C with the flow at node j increased by Δq_j . If the induced flow can be varied for a node, the ideal flow to be induced at the critical node can also be determined by means of a sensitivity vector of the same form. The terms of this new vector would reflect different Δq_j for the same node j .

It is desired that the flow at the most sensitive nodes with respect to the measure be most accurately determined since a bias in its estimate will cause a significant change in the system measure. It must be noted that the maximum improvement that can be obtained with new data is limited by the accuracy of the metering system.

Identification of Loading Condition. The identification of an ideal network loading condition for the new measurements can be developed by means of sensitivity vectors. However, an ideal load may not be possible in the real system since it may rely on consumer demands which are to some extent uncontrollable. These demands can not be arbitrarily imposed on a real time basis. The problem is more complex due to the fact that although being distributed throughout the length of the pipe, for modeling purposes the demands are lumped at the nodes. Therefore, they can not be physically measured and their estimates rely on statistics developed on a regional and demographic basis. In very few cases, when the node represents a major user or a line diversion, it can be directly measured and their demands accurately determined.

The identification of the ideal measurement load in the practical sense requires a simplification and some assumptions. In this application the measurement load will be selected from a series of pre-defined network loading conditions. These loadings are the result of forecasted

demands for different scenarios expected to occur during a normal or emergency operation. They represent potential consumer usage that can be induced to the system at the measurement time. The nodal pressures will be obtained by KYPIPE using the current knowledge of the C's for each loading condition.

Considering the dynamic behavior of consumer usage, the best approximation of the selected measurement load will be reproduced in the field during the collection period. If it is not possible to induce exactly the forecasted demands, at least the experiment can be conducted in a situation that closely approximates the desired loading condition.

Criteria to Select the Measurement Load. The selection of the ideal load can be based on two criteria. The first considers the robustness of the new measurement load. This robustness can be quantified by looking at the sensitivity of the assessment measure to small variations in the nodal demands from that load. The measure sensitivity reflects the impact of possible measurements errors and the effects of the differences between predicted and actual nodal demands.

The second approach is based on the improvement observed in the assessment measure when it is computed with the parameter estimates obtained if the new proposed measurement load is used as available information for the estimation procedure. In this case, selection of the

measurement load can proceed in the following steps after the calibration objectives have been defined:

(1) Define a series of potential network loadings for taking new measurements. The group of all possible such loads is represented by

$$\underline{Q}_P = [\underline{Q}_{p1}, \underline{Q}_{p2}, \dots, \underline{Q}_{pi}, \dots, \underline{Q}_{pL}]$$

where L stands for the number of potential measurement loads. This step corresponds to block 4 in Figure 4.2.

(2) Select a measurement load \underline{Q}_{pi} from \underline{Q}_P ;

(3) Augment the available information \underline{Q}_A by

incorporating the new proposed measurement load \underline{Q}_{pi} . The pressures corresponding to the new load, H_c , are determined by KYPIPE runs using the demands from \underline{Q}_{pi} and the current values of C's as obtained in block 10. This step corresponds to blocks 5a and 5b in Figure 4.3. The available information represents the collection of all measurement loads accessible for the estimation procedure

$$\underline{Q}_A = [\underline{Q}_{a1}, \underline{Q}_{a2}, \dots, \underline{Q}_{aM}]$$

where M is the number of available measurement loads.

The augmented measurement set is represented by

$$\underline{Q}_{Aug_i} = [\underline{Q}_A, \underline{Q}_{pi}] ;$$

The uncertainty associated with the potential measurement load Q_{pi} is assigned based on the predicted errors that may occur as a function of the accuracy of the instruments. It may be suggested the use of more accurate meters to reduce such uncertainty.

(4) Evaluate the available information and obtain pseudo-measurements if necessary using $Q_{Aug,i}$;

(5) Estimate the unknown parameters (block 5c) and find $M_{aug,i}$ (block 5d), which represents the value of the assessment measure obtained by adding the potential measurement load Q_{pi} ;

(6) Compute the improvements in the assessment measure M which was obtained using information from Q_a (block 5e)

$$\Delta M_i = M - M_{aug,i}$$

It is expected ΔM_i to be a positive number. If it is negative, the load Q_{pi} must be rejected (block 5i) since it is not contributing to improve the measure M ;

(7) Repeat steps (1) to (6) considering all loads from Q_P ;

(8) Assemble the sensitivity vector S of the form

$$S = [\Delta M_1, \Delta M_2, \dots, \Delta M_i, \dots, \Delta M_L]$$

(9) Select the potential load Q_{pi} corresponding to the largest ΔM_i as the new measurement load. The load Q_{pi} is

the desired field condition to be present when conducting the data collection. It is the load which contributes the most for improving the calibration objective, expressed by the assessment measure (See block 5k).

(10) If there is no potential load that contributes to improve the assessment measure (block 5l), the modeler has two options. He may try to improve estimates of C for individual pipes (procedure to be described later in this chapter) or he may accept the model as it is at least knowing the uncertainty expected in the results.

Identification of Critical Nodes. After selecting the network demand for the new measurements, it is necessary to identify critical nodes based on their importance in achieving the calibration objectives. The demand at the nodes should be accurately metered by installing more precise instruments or carefully determined by conducting a detailed field investigation. The critical nodes can be found using the information provided by sensitivity vector B. This vector uses the trace of the covariance matrix of the predicted pressures as the assessment measure. Other measures that could be adopted depending on the calibration objective are: the maximum nodal variance; the average nodal variance or the total system variance divided by the maximum system headloss. The procedure to find the critical nodes for the selected measurement would perform the following steps:

(1) Preselect the new measurement load. This load is represented by a vector containing the demands at individual nodes, where N_j is the number of internal nodes in the network

$$\underline{Q}_{p_i} = [q_{pi1}, q_{pi2}, \dots, q_{piN_j}]$$

Note that the above representation considers measurements on the nodal demands only. If the nodal pressures are also included, the measurement vector (MV) would take the form

$$\underline{MV}_i = [Q_{pi}, H_{pi1}, H_{pi2}, \dots, H_{piN_j}]$$

The remainder of this procedure will assume that only demands are considered;

(2) Perform steps (3) to (7) below varying j from 1 to N_j ;

(3) Vary $q_{p_{ij}}$ by a small increment Δq_{ij} . The increment will be added to the demand at node j for the measurement load i . This increment should be sufficiently large to cause changes in the assessment measure. The perturbed potential measurement load takes the form

$$\underline{Q}_{pp_i} = [q_{p_{i1}}, q_{p_{i2}}, \dots, (q_{p_{ij}} + \Delta q_{ij}), \dots, q_{p_{iN_j}}] ;$$

(4) Augment the available information by considering the perturbed load from (3). This step is preceded by running KYPIPE with the current knowledge of the C's and the load Q_{pp_i} to obtain corresponding H's;

(5) Estimate parameter C's;

(6) Perform the calibration assessment. Compute the assessment measure $M_{P_{ij}}$;

(7) Compute the sensitivity term

$$\frac{\Delta M_{ij}}{\Delta q_{ij}} = \frac{M - M_{P_{ij}}}{\Delta q_{ij}};$$

(8) Select the node j as the most sensitive node corresponding to the largest term from (7). The sensitive nodes are the ones whose measurements affect the assessment measure the most. They are critical since an error in their measurements can propagate more intensively to the simulation results.

The above procedure could be extended to determine the best distribution of demands to be induced at nodes, if more than one demand can be simultaneously considered.

Once identified, the critical nodes must be checked to determine if they carry a real localized demand. If not, it would be desirable to incorporate a new measurement load with fire conditions at the critical nodes. This would allow the demand at that node to be directly measured. The same procedure could be used to identify critical nodes considering measurements of the pressure heads.

In summary, this data collections procedure will assist in answering the following questions:

.Where to induce high demands to simulate fire flows?

.How much flow should be induced?

.Where to measure flows?

The next section introduces an alternative approach to improve simulation results based on improvements in the knowledge of the roughness factor of pipes.

Improving Knowledge of Individual Pipes

A second procedure to improve the quality of the simulation results is by directly measuring individual pipe roughness coefficients which will reduce the uncertainty in the parameter related to that link. Higher certainty in parameters from major pipes, or groups of pipes, will improve the accuracy of the model predictions. This section investigates the impact of improving the knowledge of C for individual pipes or group of pipes and the effects of these improvements in the simulation results and in the measure of the calibration assessment. The magnitude of these improvements will guide the process of selecting critical pipes to conduct individual study.

The gradients proposed below will be used to identify the pipe or group of pipes in the network that have a large contribution to the uncertainty of the simulated nodal pressures. Two types of gradients can be developed depending to the calibration objectives. They relate to the assessment of the simulated pressure at individual nodes (Gradient 1) or to the total system variance (Gradient 2). The gradient terms are defined as

Gradient 1

$$\begin{bmatrix} \frac{\partial(\text{var } Ha_1)}{\partial(\text{var } C_1)} & \frac{\partial(\text{var } Ha_1)}{\partial(\text{var } C_2)} & \dots & \frac{\partial(\text{var } Ha_1)}{\partial(\text{var } C_{N_p})} \\ \frac{\partial(\text{var } Ha_2)}{\partial(\text{var } C_1)} & \frac{\partial(\text{var } Ha_2)}{\partial(\text{var } C_2)} & \dots & \frac{\partial(\text{var } Ha_2)}{\partial(\text{var } C_{N_p})} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial(\text{var } Ha_{N_j})}{\partial(\text{var } C_1)} & \frac{\partial(\text{var } Ha_{N_j})}{\partial(\text{var } C_2)} & \dots & \frac{\partial(\text{var } Ha_{N_j})}{\partial(\text{var } C_{N_p})} \end{bmatrix}$$

Gradient 2

$$\left[\frac{\partial \text{trace}(\text{cov } H_a)}{\partial(\text{var } C_1)} \quad \frac{\partial \text{trace}(\text{cov } H_a)}{\partial(\text{var } C_2)} \quad \dots \quad \frac{\partial \text{trace}(\text{cov } H_a)}{\partial(\text{var } C_{N_p})} \right]$$

A vector formed by a column of Gradient 1 indicates which nodes are most affected by errors in the estimate of the roughness coefficient of a particular pipe. The vector formed by the transpose of each row of Gradient 1 can be used to quantify the sensitivity of the variance of the simulated pressures at a particular node with respect to each pipe roughness coefficient.

Pipes having large gradient terms are critical links since a bias in their C value causes large changes in the uncertainty measure. It is desired to improve as much as possible the estimate of C for these pipes. This improvement can be obtained by conducting individual pipe field tests. The tests would induce flows at hydrants along the pipe line to determine the roughness coefficient more accurately. Once determined, these parameters can be taken

as constant in the estimation procedure reducing the number of unknowns. This will cause improvement in the estimates of the remaining parameters since it increases the level of system overdetermination.

Gradient 1 helps identify the pipes to be selected for conducting individual field test to improve their estimates of C . It is also desirable to assess beforehand the magnitude of the improvements to be obtained with the field tests, so that judgements can be made as to the worth of the effort of conducting the actual field test. A sensitivity matrix showing changes in the trace of $\text{cov}(H_a)$, or in the $\text{var}(H_a)$, as a function of the roughness coefficient of individual pipes can be used to build sensitivity curves.

The sensitivity curves can be developed for individual pipes or for a group of pipes and would plot the trace of $\text{cov}(H_a)$ (or the $\text{var}(H_a)$ if the measure relates to individual nodes) versus the coefficient of variation of C . Sensitivity curves for individual pipes would be obtained for discrete values of the coefficient of variation (CV) of C on individual pipes equal to 1, 5 and 10%. Figure 4.4 shows the estimated form of these curves considering improvements for individual pipes and a CV equal to 10% for the remaining pipes. The x-axis reflects changes in the uncertainty level of the particular pipe considering the remaining pipes having the same CV.

Field Observations and Pseudo-Measurements

If new measurements are successfully identified in block 6, the procedure follows its main stream as shown in Figure 4.2. Measurements are taken (block 7) and pseudo-measurements are estimated (block 8). The pseudo-measurements are the additional data required for the estimation procedure and are the unmeasured pressures in block 7.

To obtain pseudo-measurements, a linear or quadratic interpolation could be used to estimate the missing H's considering the system is underdetermined. Another option would be to use non-linear regression. In this case the unknown H's (and maybe Q's) would be estimated at the same time as the C's by solving a non-linear optimization problem as proposed by Lansey and Basnet (1991). The optimization approach would use all information from the measured loads

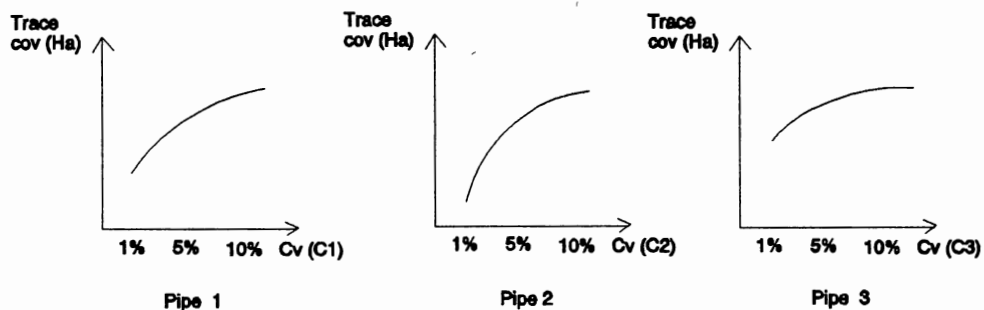


Figure 4.4. Sensitivity Curves for Pipes

and the data collected in block 6. The current knowledge of the C's (as obtained in the last iteration from block 10) could be used as the starting condition for the calibration algorithm. The stability of the nonlinear regression improves with the level of system overdetermination. One way to improve this level is by parameterization.

In block 9 the measurements are augmented based on assumed network demands and simulated pressures. This step will only be required in the first stages of the calibration when the information available is not enough to guarantee system observability.

Parameter Estimation

Two major parameters are considered to model a water distribution system as mentioned in Chapter II. They are the pipe roughness coefficient and the consumer demands. The procedure to estimate the unknown parameters in a network system is dependent on the type and number of the unknowns, their distribution throughout the system and the amount and characteristics of the input data available.

For known values of nodal pressure heads and using the Hazen-Williams equations to express head loss terms, the flow rate in the pipe appears as a linear function of the roughness coefficient C (See equation 2.10). By expressing the system equations in terms of the nodal equations, one can write N_i independent continuity equations where N_i is the number of internal nodes in the network exclusive of the

fixed grade node or the supply node.

When pipe roughnesses are unknown and considering a single demand, the number of the system equations is always smaller than the number of the pipes in the network (See Eq. 2.4). Since these equations are being solved for the roughness factors, the solution is said to be underdetermined. Two options can be applied to overcome this problem. The first consists of reducing the number of unknowns by assigning the same C for a group of similar pipes (network parameterization). The second option uses additional input data from other demand conditions. Each new condition allows one to write N_d new equations while the number of unknown parameters, N_p , remains the same.

The use of information from multiple demands was considered in this research for the development of the parameter estimation procedure. The ability to consider network parameterization was also incorporated in the program developed and its benefits were investigated in Chapter VI as an alternative option to reduce computational effort.

A linear regression approach was developed to estimate parameters as described below. Next, the process of handling uncertainties in input data is explained. The last sub-section introduces a procedure to transfer these uncertainties to the parameter estimates. The parameter estimation procedure is shown in blocks 10 and 11 in Figure 4.2 and is expanded in Figure 4.5.

Linear Regression for Coefficient C

Considering the pressure heads and external demands in the network as known values, the estimation of unknown parameters C can be obtained by solving a system of linear equations of the form

$$Q = X C \quad (4.2)$$

where

Q = $N_T \times 1$ vector of external nodal demands for all loads

X = $N_T \times N_p$ matrix containing information regarding network physical characteristics, geometry and pressures. Their elements are the first term of Eq. (2.10)

C = $N_p \times 1$ vector of unknown parameters

N_p = number of parameters to be estimated

N_T = $N_J * N_{Load}$

N_{Load} = Number of measurement loads.

The unknown roughness factors C can be estimated by the least squares method (Haan, 1977) as

$$\hat{C} = (X^t X)^{-1} X^t Q \quad (4.3)$$

where

\hat{C} = $N_p \times 1$ vector of estimates of parameter C

X^t = Transpose of X

$()^{-1}$ = inverse of the matrix inside parentheses.

The procedure requires the inversion of a N_p by N_p matrix. The multiple linear regression corresponds to block 10b in Figure 4.5.

The regression approach requires knowledge of the pressures at all nodes and demands at most of the points in the system. It is possible to estimate C's for a limited number of unknown Q's. However, for each unknown Q one equation is lost for each load. If a pressure is unknown at a particular node, a number of equations equal to the number of pipes connecting to that node is also lost. Since \underline{X} is a sparse matrix, care must be taken to keep the rank of the $(\underline{X}^t \underline{X})$ matrix equal to N_p after dropping the equations corresponding to the unknown Q's and H's.

When some of the H's in the network are not known, a nonlinear optimization approach is available from the work of Lansey and Basnet (1991). The results, however, will be limited by the type and quantity of input data which will determine the system observability level. As noticed earlier the pressure head at unmonitored locations can be found using this procedure or another interpolation scheme.

Uncertainty in Calibration Data

Uncertainties in input data result from several sources. They can be caused by the accuracy of the metering device, the adequacy of the meter scale for the range being measured, and other aspects in the data collection and acquisition process (Simmonds and Laverty, 1980). It can

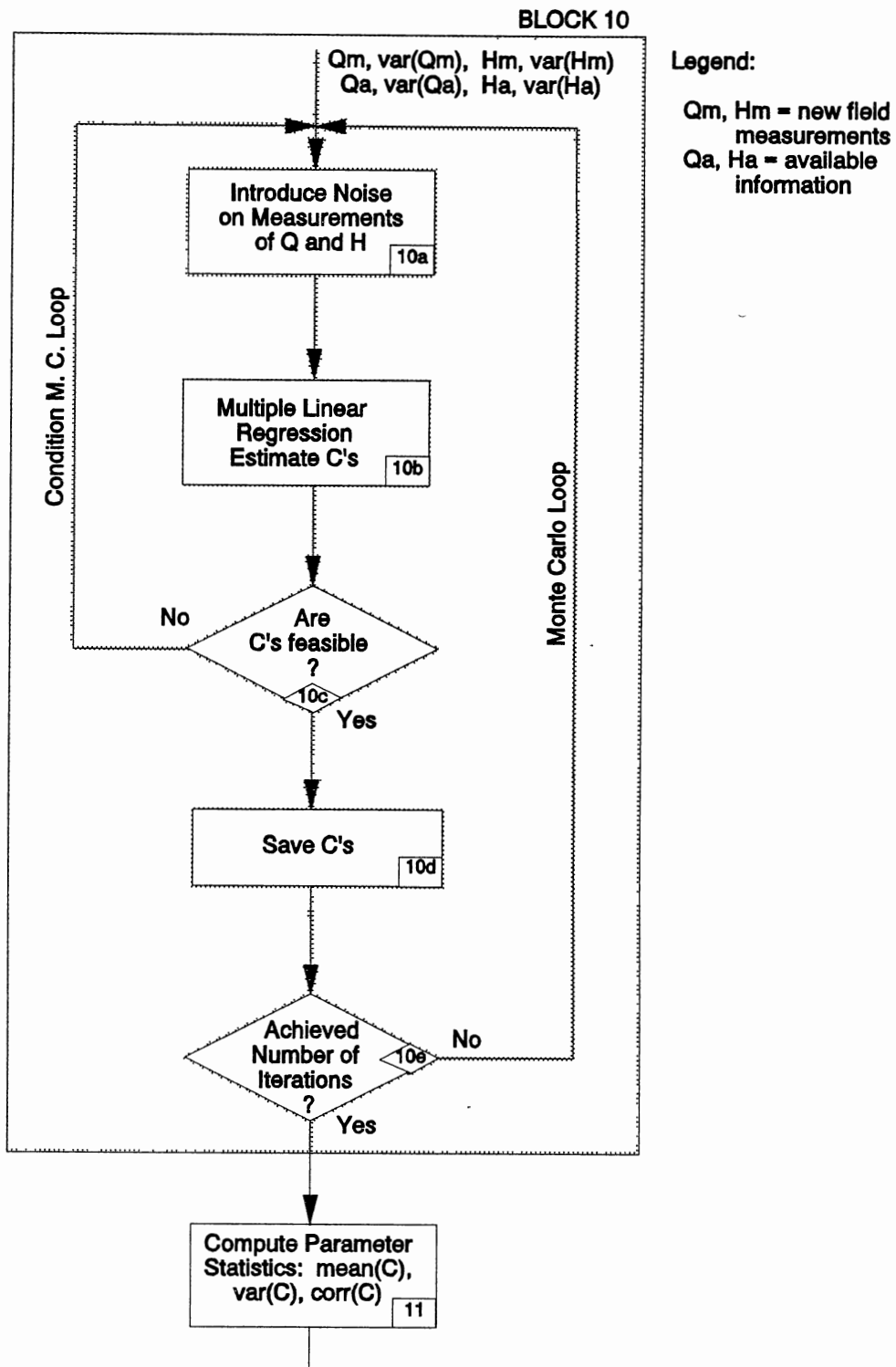


Figure 4.5. Parameter Estimation Process Diagram

also reflect the criteria used to estimate the consumer demands. To model these uncertainties an error term was introduced to corrupt the mean values of the pressures and nodal demands.

The errors associated with input data were assumed normally distributed. To create the vector of perturbed demands, standard normal deviates were generated individually for each node from all loads. The magnitude of the error term was adjusted to reflect the prescribed level of confidence in estimating Q for that node, expressed by means of its coefficient of variation (CV). If E is the normal deviate, the perturbed demand is computed for each node as

$$qp_{ij} = q_{ij} + [E * \sigma_{q_{ij}}] \quad (4.4)$$

where

- qp_{ij} = q perturbed at node i , for load j
- q_{ij} = mean q at node i , for load j
- E = standard normal deviate
- $\sigma_{q_{ij}}$ = $CV_{ij} * q_{ij}$, is the standard deviation of q_{ij}
- CV_{ij} = coefficient of variation of q_{ij} .

The procedure of generating errors for pressure heads is similar to the one for the demands. The difference lies in the way of assigning the uncertainty levels for each pressure. In this case the level is specified by means of the standard deviation of the H 's instead of the CV. The

perturbed pressure heads are computed as

$$HP_{ij} = H_{ij} + [E * \sigma_{Hij}] \quad (4.5)$$

where

- HP_{ij} = perturbed pressure at node i, for load j
 H_{ij} = mean pressure at node i, for load j
 E = standard normal deviate
 σ_{Hij} = standard deviation of H_{ij} .

Uncertainty in Estimated Parameters

The procedure to find the uncertainty in the estimates of C requires the solution of the regression problem many times within a Monte Carlo approach. At each time, random deviates are introduced in the input variables according to the assumed distribution (See block 10a, Figure 4.5). A detailed description of Monte Carlo methods can be found in Beck and Arnold (1977) and Ang and Tang (1984). Before being accepted as valid results, the regressed C's are checked against feasibility limits, as shown in block 10c. If the regressed C's are not feasible, the series of input errors is rejected and the process is repeated again. This procedure is represented by the conditioned Monte Carlo loop in Figure 4.5. At each run the regressed parameters are saved (block 10d) and their statistics are computed at the end of the process (block 11). They refer to the mean; standard deviation; and covariance, correlation and factor

loading matrices of the C's. Parameter uncertainty is quantified by the variances of C's.

Because the parameter C is highly sensitive to changes in pressures and external flows, the process may require intensive computational efforts. However, once completed, it provides a good approximation to the mean and variance of the parameters and gives an idea of their distribution.

A second way of computing the variance of the estimates of C's is directly from the regression performed in Block 10 by means of the standard error of the estimates. This can only be done if the number of measured loads available to the parameter estimation process is large enough to provide enough information to assure an adequate level of system overdetermination. This condition, although theoretically possible, is constrained by the few data available due to the high costs involved with the current practices and available technology for data collection and acquisition systems.

A third method for estimating the variance is to use first order approximation which is similar to the assessment procedure. For this highly nonlinear system these estimates were quite poor, particularly for individual nodes. However, the relative magnitudes were reasonable between nodes and it may be used to save computational time during early iterations of the calibration process.

At this point the calibration loop is closed and a new iteration is completed. It is expected that the parameters

improve at each new iteration. The improvements are verified when the C's converge to the correct values and their uncertainties decrease. Since the C's are very sensitive to the measurements, changes in the measurements may cause large variability in the regressed values. An alternative to decrease such sensitivity is to reduce the number of unknown parameters by parameterizing the network.

Network parameterization contributes to increase the level of system overdetermination and improve the regression results. Parameterization, however, may introduce modeling errors although improving parameter estimates. The trade-off between the effects of both types of errors has been studied by Yeh and Yoon (1981) for aquifer systems and should be extended to the water distribution systems.

It may be found for the network under study that it is not possible to meet the pressure requirements at all nodes. If this happens a question may arise whether the calibration process should stop or more data should be collected. A design revision may solve the problem (if there is one) so to insure good estimates of the mean C and to guarantee an accurate model as a basis for design decisions, the calibration procedure should be continued.

CHAPTER V

APPLICATION OF PARAMETER ESTIMATION AND ASSESSMENT PROCEDURES

The purpose of this chapter is to present the methodology for parameter estimation and assessment introduced in Chapter IV. A simple network system containing 11 pipes is used to illustrate the several steps involved with the process and its input requirements and assumptions.

The objectives of the application developed in this chapter are to: (1) apply the methodology to a small network to demonstrate the model's capability for handling uncertainty in input data and parameters, (2) show the usefulness and consistency of the measure adopted to quantify the uncertainties in the results derived from the model, (3) demonstrate the impact of different levels of uncertainty in calibration data and its effects in the model prediction uncertainty, and (4) compare different methods to assess the prediction ability of the model according to the goals set for the calibration.

To accomplish the above objectives, the chapter is divided into four sections. The first section describes the physical characteristics of the distribution network used to

conduct experiments for the application. The second section presents the demand loads used as available information for the estimation of the unknown roughness parameters. This information consists of measurements and pseudo-measurements of pressure heads and of estimation of consumer demands. In the third section, the parameter estimation section, the assumptions made to model the uncertainties are given and how these uncertainties are transferred to the parameters is discussed. The last section assesses the model prediction uncertainty by the two methods described earlier in Chapter IV; Monte Carlo simulation and first order approximation of the variance.

Network Description

The water distribution system used to run experiments for this application is a hypothetical network formed by 11 pipes, 7 internal nodes and 4 loops, and is called Network 1. Figure 5.1 shows network characteristics and numbering scheme. The source of water is at node 8 which has a fixed grade at elevation 200 feet. The network is placed in a flat area with elevation of 100 feet at all nodes.

Water is delivered to the network by gravity through a 1500 ft, 22-inch pipe line corresponding to link 11. The total pipe length is 6200 feet. Consumer withdrawals from the system total 9 cubic feet per second (cfs) under normal conditions.

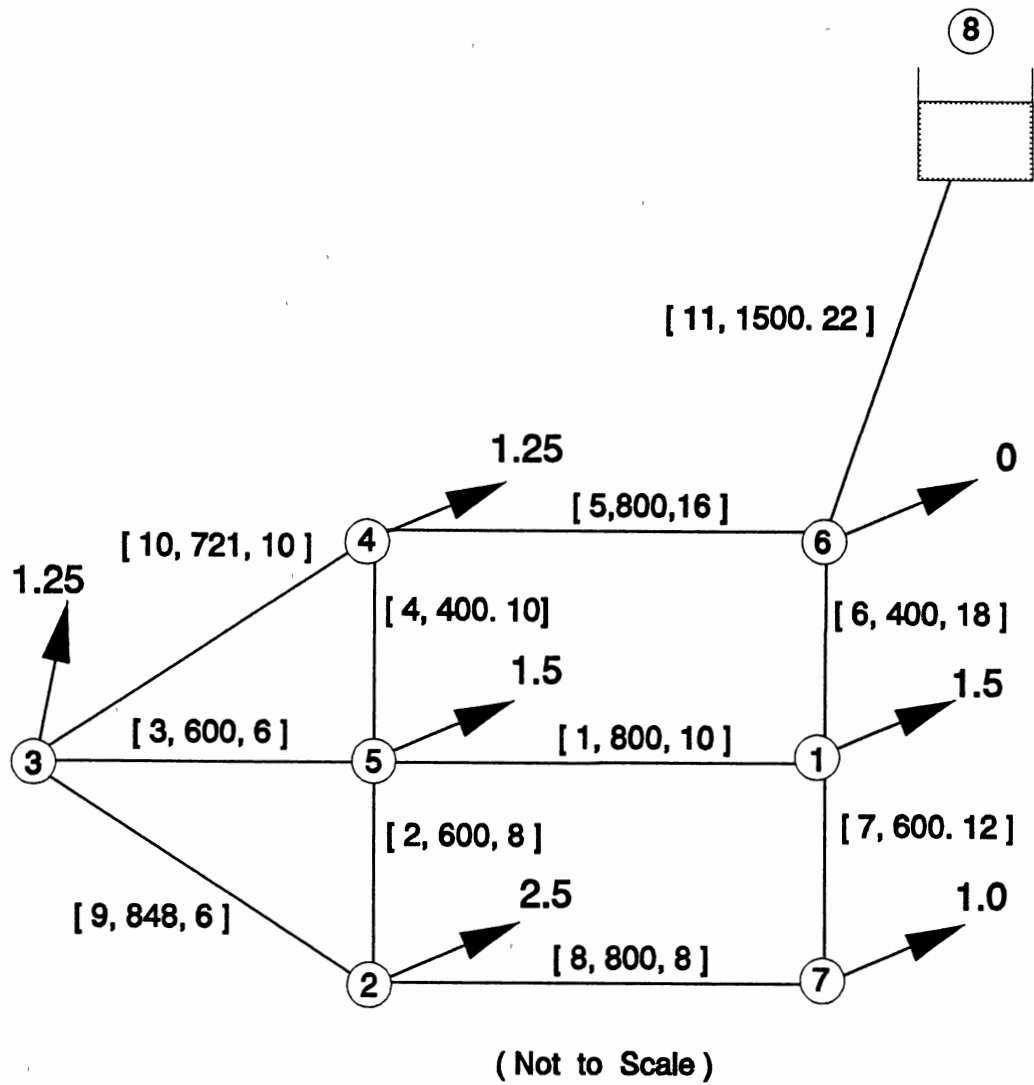


Figure 5.1. Geometry and Numbering Scheme-Network 1

A roughness factor for the pipes (Hazen-Williams coefficient C) equal to 100 was assigned to all links in the system. These values, considered as the true C's, are the unknown parameters to be estimated. They were not provided to the model at any time. The true C's were used to simulate measurements of nodal pressure. These pressures are used as input variables in the estimation procedure. Pipe characteristics for Network 1 are shown in Table 5.1.

Considering the nodal pressures and demands as known for a loading condition, the system configuration allows one to write 7 mass balance equations (one for each internal node). The unknowns are the 11 pipe roughness coefficients. So, having information from only one demand pattern makes the system underdetermined because there are 11 unknowns and only 7 equations. To overcome the problem of underdetermination, multiple demand patterns were considered as described in the following section.

Demand Patterns

The input data necessary to allow the estimation of the unknown roughness parameters consists of nodal pressure head measurements and consumer demand estimates. Measurements or pseudo-measurements of pressure head were considered known for all nodes for each load. In a real scenario, these values are measured during the data collection efforts at selected points in the system. The pseudo-measurements are interpolated or estimated pressures that were not directly

TABLE 5.1
PIPE CHARACTERISTICS FOR NETWORK 1

Pipe # (1)	Diameter (in) (2)	Length (ft) (3)
1	10	800
2	8	600
3	6	600
4	10	400
5	16	800
6	18	400
7	12	600
8	8	800
9	6	721
10	10	848
11	22	1500

measured.

The other input data are the external nodal demands. External demands are usually estimated based on geographic location and consumer usage. They may be measured when representing a localized withdrawal or the consumption of a large water user. In the present application demands were adopted and pressures were simulated using the true value of the roughness parameters. The network simulator KYPIPE (Wood, 1981) was used to generate the necessary input data for the estimation procedure.

Input information to estimate parameters for network 1

was considered for five loading conditions. The loading conditions available represented consumer withdrawals under the following usage patterns: normal (N), peak (P), fire at node 2 (F2), fire at node 3 (F3), and fire at node 5 (F5). The normal load corresponds to the average daily demand. Other demands reflect different operating conditions. The peak demand has considered nodal consumptions which are twice as high as the normal external flows. Fire fighting conditions were simulated at nodes 2, 3, and 5 which required flows equal to 5.5, 6.0 and 6.5 cfs, respectively. During this period, consumer withdrawals at remaining nodes were reduced proportionally to their normal values. Table 5.2 shows demands and corresponding pressure for all nodes for each loading condition.

The demands and pressures were used as mean values. Random errors were added to reflect uncertainties due to the data collection and processing and due to the estimation of pseudo-measurements. A technique to model these uncertainties is discussed in the next section.

Estimation of Model Parameters with Uncertainty Data

The structure of the error term to perturb the mean values of the input variables was assumed known. The errors introduced in the estimated consumer demands were considered normally distributed with zero mean. Their magnitude was adjusted by changing their standard deviation to reflect the

TABLE 5.2
 AVAILABLE INFORMATION USED TO ESTIMATE
 PARAMETERS FOR NETWORK 1

Loading Condition (1)	Node Number (2)	Demand (cfs) (3)	Pressure (ft) (4)
Normal	1	1.50	194.72
	2	2.50	186.82
	3	1.25	190.20
	4	1.25	193.22
	5	1.50	191.43
	6	0.00	195.71
	7	1.00	192.62
	8	-9.00	200.00
Peak	1	3.00	180.94
	2	5.00	152.42
	3	2.50	164.64
	4	2.50	175.52
	5	3.00	169.06
	6	0.00	184.52
	7	2.00	173.35
	8	-18.50	200.00
Fire at node 2	1	1.00	193.29
	2	5.50	167.85
	3	0.80	187.64
	4	0.90	191.61
	5	1.10	188.38
	6	0.00	194.59
	7	0.90	189.02
	8	-10.20	200.00
Fire at node 3	1	0.75	194.18
	2	1.25	185.35
	3	6.00	169.08
	4	0.63	190.26
	5	0.75	189.17
	6	0.00	194.91
	7	0.50	192.71
	8	-9.88	200.00

TABLE 5.2 (Continued)

Loading Condition (1)	Node Number (2)	Demand (cfs) (3)	Pressure (ft) (4)
Fire at node 5	1	0.80	191.66
	2	1.75	179.69
	3	0.90	183.29
	4	1.00	187.86
	5	6.50	179.68
	6	0.00	192.97
	7	0.80	189.31
	8	-11.75	200.00

level of uncertainty in the estimation of the nodal demands.

The coefficient of variation of the demands, CV of Q's, was used to assign the uncertainty level to be modeled at each node. This coefficient corresponds to a normalized standard deviation. It is computed by dividing the standard deviation of Q by its mean value. By using the coefficient of variation, the standard deviations of the consumer demands are scaled to the mean flow at each node.

The Fortran program developed, as documented in Appendix D, has options to consider uncertainty in Q's by assigning: (a) the CV's individually for each node, (b) the same CV's for the nodes from the same load, or (c) the same CV for all nodes for all loads. If the last two options are used, higher uncertainties will be assigned to the nodes

with larger withdrawals. The program has another option to limit the maximum standard deviation at the nodes to prevent unreasonably high values for the errors especially at the nodes carrying fire demands.

The uncertainty in measurements of nodal pressures were simulated by the same process as described above. The error term was also assumed to be normally distributed. The only difference from the above procedure is that the standard deviation of the H's were used to assign the uncertainty levels. Three options have been provided in the program for accepting the uncertainty for the H's. The nodal standard deviation can be assigned: (a) individually for each node, (b) grouped for the nodes from the same load, or (c) a single value for all nodes from the several loading conditions.

This application attempts to estimate roughness coefficients for each pipe given the uncertainties in the demands used as input data. Different levels of uncertainty for the Q's were considered to show their effects in the parameter estimate uncertainty. Coefficients of variation of the Q's equal to 1%, 5%, and 10% were considered. The pressures at all nodes were assumed to be known with certainty. No error was added to their mean values. The uncertainty in both demands and pressures was incorporated in the application in Chapter VI.

Two criteria were used for conditioning input data and the results of the regressions. The first limits the sigma

bounds of the normal distribution used to generate random errors to perturb the Q's. The sigma bounds are the number of standard deviations from the mean used to accept the error term which is applied to the nodal demands. If the randomly generated error falls within plus or minus n standard deviations it is accepted as a valid term. In this case the Q's perturbed will be within the n sigma bounds of the distribution (Haan, 1977, pg. 87). Choosing n equal to 1 corresponds to 68% of the normal distribution. Most of the runs performed assigned n equal 3 corresponding to 99.72% of the distribution.

A second conditioning criterion was used to determine the acceptability of the errors introduced considering a valid solution for the system equations. This criterion introduces physical reasoning to accept the randomly generated errors added to the input data. The criterion establishes limits for the regressed parameter C's, according to the pipe material, diameter and age. If a regressed C has a value outside of the acceptable range, the regression is rejected and a new set of errors is generated to perturb all measurements.

Statistical analysis for each pipe roughness coefficient was computed based on 500 regressions for parameter C's within the Monte Carlo approach. The number of regressions required to assure convergence of the mean and standard deviation of the C's is investigated in Chapter VI.

Table 5.3 shows estimates of C's for selected pipes. Appendix A contains results for all pipes. Column 1 of Table 5.3 shows the coefficient of variation of the Q's used to perturb the mean value of that variable. Column 2 states the type of statistics performed as the mean C or the standard deviation (std) of the C's. Columns 3 to 8 show statistics for selected pipes. The trace of the covariance matrix of the regressed C's appears in column 9 as a measure of the total system variance.

Values from Table 5.3 reveal that as the CV of Q increased so did the variability in C's and the magnitude of the trace of $\text{cov}(C)$. The trace of the covariance matrix of the C's varied from 8.7 for CV of Q equal 1% to 888 for CV of Q equal to 10%. The standard deviation of C for pipe 3 increased from 1.5 units to 15.6 units for CV of Q varying from 1% to 10%. For pipe 9 the variability in C changed from 1.4 to 13.2 units for the same conditions.

Pipes 3 and 9 are considered weak pipes in the system. Since they carry low flow, the head loss through their links are small and a wider range of the C values does not significantly affect the pressure distribution through the system. The opposite occurs for pipes carrying large flows such as pipes 5 and 11. For these pipes, a small variation in their roughness can cause significant changes in the head loss affecting pressures at all nodes influenced by these links. Average standard deviation of C for pipe 11, as shown in column 8 from Table 5.3, is only 2.2 units for CV

TABLE 5.3
ESTIMATED PIPE ROUGHNESS C FOR DIFFERENT
LEVELS OF UNCERTAINTIES IN Q'S

CV Q (1)	Statistics (2)	Pipe #						Trace cov(C) (9)
		1 (3)	3 (4)	5 (5)	8 (6)	9 (7)	11 (8)	
1%	mean C	99.9	100.0	99.9	99.9	100.0	99.9	8.7
	std. C	0.9	1.5	0.4	1.0	1.4	0.2	
5%	mean C	99.9	99.9	99.9	99.8	100.2	99.9	223
	std. C	4.3	7.7	2.0	5.2	7.1	1.1	
10%	mean C	100.6	101.3	99.5	99.8	100.0	99.9	888
	std. C	9.1	15.6	4.1	10.0	13.2	2.2	
TRUE C		100.0	100.0	100.0	100.0	100.0	100.0	

of Q equal to 10%. This means that there is less uncertainty in the C for that pipe as compared with those from pipes 3 and 9.

It should be reiterated that the C's shown in Table 5.3 are means based on 500 individual estimates. Individual estimates may deviate from the mean shown with the probability of a large deviation increasing as the std of C increases.

Correlation among parameters was investigated based on data from 500 regressions used to compute parameter statistics. The correlation matrix of the C's for the

calibration run considering CV of Q equal 5% is listed in Table 5.4. It should be emphasized that this correlation has a statistical base and results from the effects of input data uncertainty. The errors introduced in the demand and pressure of a particular node cause proportional effects in the flow at pipes connected to that node, which influences the estimation of their parameters. The high values observed in Table 5.4 stress the importance to preserve the correlation among parameters when transferring parameter uncertainty to model outputs.

The statistical correlation should be differentiated from physical correlation that may develop over a long period of time as roughness factors change due to incrustation in the pipes.

In this application, the calibration accuracy can be directly assessed by comparing the estimates of C with their true values. However, this cannot be done in real applications nor does it provide any idea how well the calibrated model can predict pressures for other demand patterns. One way to overcome this problem is by looking at the state variable, H_a , the pressure head computed by the model using the knowledge of C's obtained with different calibration levels. This step is discussed in the next section.

Calibration Assessment

The objective of calibration assessment is to find the uncertainty in the computed nodal pressure head based on the uncertainty of calibrated parameter C's. Critical demands will be examined in this stage to judge the accuracy of the calibrated model. The modeler can then decide if the calibration is acceptable for the intended application. In the present application, the assessment was performed with parameter C's obtained with levels of uncertainties reflecting CV of Q's equal to 1%, 5%, and 10%.

The demand selected to assess the calibration reproduces the flow requirements to fight a fire located at node 2. The magnitude of the flow at that node is 7.5 cubic feet per second. The remaining nodes in the distribution network had their demands reduced to 80% of the normal values. Table 5.5 shows nodal demands for the assessment load.

The calibration assessment was performed by Monte Carlo simulations and by first order estimation of the variance. Both methods considered correlation among parameters and the uncertainties associated with each parameter. As mentioned in Chapter IV, the Monte Carlo method generates multi-variate C variables to preserve the parameter correlation found in the estimation module, as shown in Table 5.4. The procedure to find correlated random variables is shown in Appendix C.

TABLE 5.4
CORRELATION AMONG PARAMETERS FOR CV OF Q EQUAL 5%

	R O U G H N E S S P A R A M E T E R										
	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈	C ₉	C ₁₀	C ₁₁
C ₁	1.000										
C ₂	0.552	1.000									
C ₃	0.517	0.093	1.000								
C ₄	-0.014	0.566	0.125	1.000							
C ₅	-0.249	0.263	0.056	0.760	1.000						
C ₆	0.596	0.139	0.096	-0.334	-0.370	1.000					
C ₇	-0.417	-0.360	-0.449	-0.191	-0.050	0.406	1.000				
C ₈	-0.456	-0.406	-0.487	-0.228	-0.052	0.325	0.932	1.000			
C ₉	-0.312	-0.123	-0.072	0.106	0.506	-0.158	0.147	0.176	1.000		
C ₁₀	-0.387	-0.190	0.009	0.165	0.697	-0.244	0.125	0.165	0.762	1.000	
C ₁₁	0.320	0.357	0.136	0.366	0.544	0.578	0.323	0.248	0.302	0.392	1.000

TABLE 5.5
ASSESSMENT LOAD FOR NETWORK 1 WITH FIRE AT NODE 2

Condition (1)	Node Number (2)	Flow (cfs) (3)
Fire at node 4	1	1.2
	2	7.5
	3	1.0
	4	1.0
	5	1.2
	6	0.0
	7	0.8
	8	-12.7

The FOSM approach accounts for the same correlation by using the full covariance matrix of the parameters, $cov(C)$, in equation 4.1. This matrix would be replaced by the $var(C)$ if correlation among parameters would not be considered.

To quantify the uncertainty in model predictions the trace of the covariance matrix of computed pressures, $trace\ cov(H_a)$, was adopted as a measure representing the total system variance. This trace is obtained by adding the diagonal elements of the matrix which corresponds to the summation of the pressure variance at each node. Results obtained by applying Monte Carlo and FOSM methods are presented and compared in the following sub-sections.

Monte Carlo Method

The Monte Carlo method requires knowledge of the multivariate distribution of the roughness parameters. C parameters were found to be normally distributed for all major pipes in the network. In this application it was assumed that all C's were normally distributed. This assumption was verified in Chapter VI.

The procedure for generating the multivariate distribution consists of drawing a random parameter C for each pipe from the normal distribution such as to preserve the mean, standard deviation and correlation among the C's. The parameters are then used in a network simulation model to compute pressure at the nodes for the selected assessment condition. The process is repeated many times to allow an approximation of the distribution of model outputs. In this way, the parameter uncertainty is transformed into model prediction uncertainty.

Table 5.6 shows model predicted pressures and their uncertainty for the assessment load. These pressures were computed using parameters from different calibration runs. Before the generated C's were passed to the network simulator they were checked for consistency. This criterion requires the range of C's to be within acceptable limits. Values from this table represent the average of five realizations. For each realization the statistics were computed based on 1000 accepted simulation results. Table

5.6 includes pressures for selected nodes. Appendix A contains results for the entire network.

The statistics presented in column 2 on Table 5.6 are the average of the mean pressures and the average of the standard deviations of the five realizations. These values show that using calibrated parameters with CV of Q equal to 10%, the average standard deviations of the pressures ranged from 0.32 feet at node 6 to 4.0 feet at node 2. Node 2 has the simulated flow of 7.5 cfs and accounts for the majority of the total system variance. The high flow in pipes carrying water to node 2 causes high head loss through their links. Consequently, variability in the C 's will affect directly the pressures at nodes receiving water from these links. Column 7 shows the trace of the covariance matrix of computed pressure. This value expresses the total system variance which is used to quantify uncertainties in the results derived by the model.

The standard deviation of pressures at individual nodes increased as the assessment was performed with parameters having more variability. Small variability was found for nodes 1 and 6, with standard deviations for CV on Q equal 10% equal to 0.4 ft and 0.32 ft respectively. These nodes have low or no external flow. They are also located close to the source so the flow reaching the node only passes through a few pipes.

TABLE 5.6
CALIBRATION ASSESSMENT FOR NETWORK 1
BY MONTE CARLO METHOD

Par. Estimation		Mean and Std of Pressures at Nodes				
CV %	Q Statistics	Node Number				Trace cov(Ha) (ft ²)
(1)	(2)	1	2	3	6	(7)
		(3)	(4)	(5)	(6)	
1	mean	189.9	145.8	180.8	191.9	0.18
	std	0.04	0.40	0.08	0.03	
5	mean	189.9	145.6	180.8	191.9	4.74
	std	0.20	2.03	0.42	0.16	
10	mean	189.9	145.2	180.8	191.9	18.44
	std	0.40	4.00	0.82	0.32	
True H		190.0	145.9	180.9	191.9	

First Order Analysis

The first order, second moment approach estimates the covariance of the predicted pressures by means of the sensitivity matrix of the model responses with respect to the parameters and by the covariance matrix of the parameters (equation 4.1). This approach is used as an efficient method to compute the covariance matrix of predicted pressures, a matrix used to define a measure of uncertainty in model outputs. The mean values of the H's

were computed by using KYPIPE and assigning the mean value of the roughness parameters. The number of network simulations required to compute the sensitivity matrix is the same as the number of pipes in the system plus one. In this example 12 network simulations were performed to compute the sensitivity matrix. This number is insignificant if compared with the 1000 simulations performed for the Monte Carlo method.

Predicted pressure heads, their uncertainties and the trace of the covariance matrix computed by FOSM method are presented in Table 5.7 for all calibration runs. The values compared well with the ones from Table 5.6 showing the Monte Carlo results. Both methods were capable of predicting pressures at all nodes within 1% of their true values. Appendix A includes pressures for all nodes.

The standard deviations computed by the FOSM approximation were slightly higher than those computed by the Monte Carlo method. The maximum difference was found for node 2 which is a critical node carrying the fire demand. The standard deviation obtained by the Monte Carlo method was 4.0 ft for CV on Q equal 10% compared to 4.13 ft from the FOSM method. The difference in the trace computed by the two methods was between 5 and 7%. The savings in computational time by using FOSM appears to be worthwhile given the relative small difference in the computed values.

The next chapter applies the calibration procedure to a larger network. The application developed investigates the

approximation of the distribution for the regressed parameters and for the model outputs. It also presents other aspects involved with the methodology.

TABLE 5.7
CALIBRATION ASSESSMENT FOR NETWORK 1
BY FIRST ORDER METHOD

Par. Estimation		Mean and std of Pressures at nodes				
CV %	Q Statistics	Node Number				Trace cov(Ha)
(1)	(2)	1	2	3	6	(7)
		(feet)				(ft ²)
		(3)	(4)	(5)	(6)	
1	mean	189.9	145.8	180.8	191.9	0.19
	std	0.04	0.41	0.08	0.03	
5	mean	189.9	145.7	180.8	191.9	5.08
	std	0.21	2.10	0.44	0.17	
10	mean	189.9	145.5	180.8	191.9	19.65
	std	0.42	4.13	0.85	0.34	
True H		190.0	145.9	180.9	191.9	

CHAPTER VI

EVALUATION OF THE ASSESSMENT CRITERIA

This chapter presents results of several applications developed to illustrate the calibration approach considering uncertainties in input variables, in parameters, and in simulation results. One of the purposes of the applications is to show the sensitivity of the measure of prediction uncertainty to different types and levels of input uncertainties. Also the applications demonstrate the usefulness of the assessment measure and how it can be used as a criteria to judge the calibration accuracy according to the intended use of the results.

The experiments conducted in this chapter illustrate how the measure of the models' prediction uncertainty is affected by: (a) different levels of uncertainty in estimation of consumer demands Q 's; (b) different levels of uncertainty in measurements and pseudo-measurements of pressure heads H 's; (c) different types of assessment demands; (d) calibration performed with measurements from a low stress network condition; (e) calibration performed with different amount of available information; and (f) calibration performed considering a parameterized network. An investigation was performed to examine the distribution

of the regressed roughness parameters and model output resulting from the distribution of the input variables. A second investigation was also conducted to find the number of simulations required for the Monte Carlo approach.

Network Description

The distribution network used to run experiments for the application is formed by 16 pipes, 12 internal nodes and 4 loops. The source of water is at node 13 which has a fixed grade at elevation 50 feet. The highest point in the system is located at elevation 165 feet. The maximum difference in elevation between this node and the lowest point is 35 feet. The water is delivered to the network through a 10000 ft, 24-inch pipe line corresponding to link 1. There is a pump in this line to provide enough energy such that the available pressure at all points in the network is within adequate limits. For modeling purposes, the effect of the pump was reproduced by adding its contributing pressure to the grade at node 13.

The network comprises 75200 feet of pipe line. The total system demand is 29.6 cubic feet per second (cfs) under normal demand conditions. Figure 6.1 shows the network characteristics and numbering scheme. Pipe length and diameter are presented in Table 6.1.

Measurements under five network loading conditions were considered to estimate the unknown pipe roughness coefficients. These loads are referred to as: Normal (N),

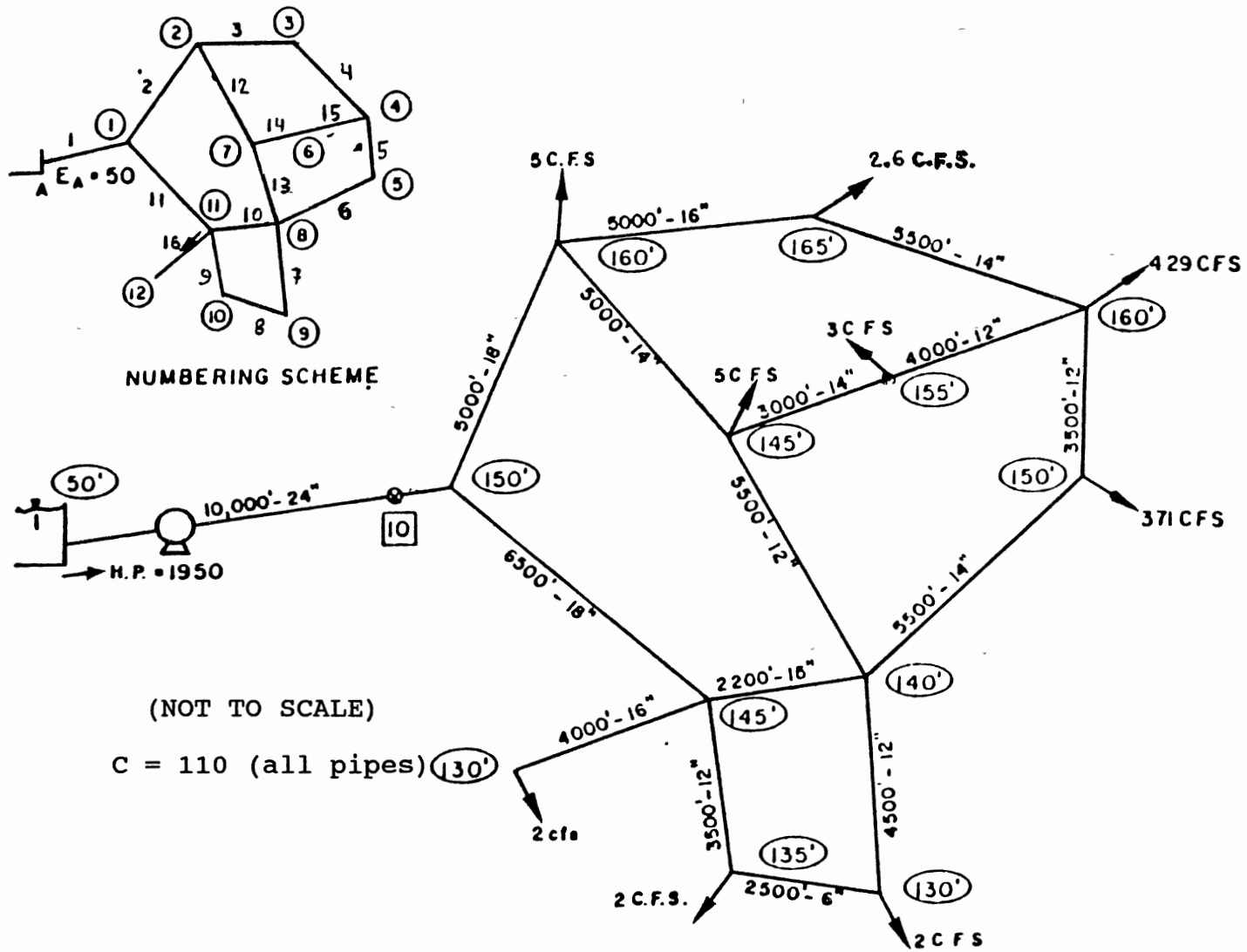


Figure 6.1 Geometry and Numbering Scheme for Network 2

TABLE 6.1
PIPE CHARACTERISTICS FOR NETWORK 2

Pipe # (1)	Diameter (in) (2)	Length (ft) (3)
1	24	10000
2	18	5000
3	16	5000
4	14	5500
5	12	3500
6	14	5500
7	12	4500
8	6	2500
9	12	3500
10	15	2200
11	18	6500
12	4	5000
13	12	5500
14	14	3000
15	12	4000
16	16	4000

Peak (P), Slack (S), Fire at node 3 (F3), and Fire at node 8 (F8). The normal load corresponds to the average daily demand. The peak demand considered normal external flows increased by 40%. The slack condition had the normal demands reduced by 60%. A fire fighting situation was simulated at node 3 with a required flow of 4.5 cfs. During this period, consumer withdrawals at remaining nodes were reduced to 80% of their normal values. A fire fighting condition with the same characteristics was reproduced at

node 8. Table 6.2 shows the demand for all nodes for each load. Pressure head and demands were made available for all nodes under each loading condition.

To generate the required input data, the demands were adopted and pressures were simulated using KYPIPE with the true value of the roughness parameters. Random errors were added to reflect uncertainties in input variables in the same manner as described in Chapter V.

Uncertainties in Nodal Demands

Several runs were performed to study different levels of uncertainties in calibration data and their effects in the parameter estimates. In the previous chapter the pressures were considered as exact values. Now, uncertainties are introduced in Q's as well as in the H's. The effects of these different sources of uncertainties in input data were individually assessed. In the first series of runs a coefficient of variation (CV) on Q's equal to 1, 5, 10, and 25% were considered. The uncertainty in the H's were kept constant with a standard deviation (std) of 0.25 ft for all nodes. In the second series of runs uncertainty levels of the H's varied while the uncertainties in the Q's remained constant. Values for the std of the H's equal to 0.1, 0.25, 0.5 and 1.0 ft were considered and a CV of Q's equal to 5% was used for all cases. The errors introduced in Q's and H's were assumed normally distributed.

Statistical analysis for each pipe roughness

TABLE 6.2
 AVAILABLE INFORMATION USED TO ESTIMATE
 PARAMETERS FOR NETWORK 2

Loading Condition (1)	Node Number (2)	Demand (cfs) (3)	Pressure (ft) (4)
Normal	1	0.00	474.76
	2	5.00	376.81
	3	2.60	348.74
	4	4.29	327.27
	5	3.71	328.89
	6	3.00	327.27
	7	5.00	335.59
	8	0.00	361.35
	9	2.00	354.66
	10	2.00	376.02
	11	0.00	391.97
	12	2.00	389.19
	13	-29.60	617.00
Peak	1	0.00	481.75
	2	7.00	299.10
	3	3.64	246.76
	4	6.01	206.72
	5	5.19	209.73
	6	4.20	206.71
	7	7.00	222.24
	8	0.00	270.28
	9	2.80	257.80
	10	2.80	297.63
	11	0.00	327.38
	12	2.80	322.18
	13	-41.44	747.00
Slack	1	0.00	561.77
	2	3.00	523.74
	3	1.56	512.84
	4	2.57	504.51
	5	2.23	505.13
	6	1.80	504.50
	7	3.00	507.74
	8	0.00	517.74
	9	1.20	515.14
	10	1.20	523.44
	11	0.00	529.63

TABLE 6.2 (Continued)

Loading Condition (1)	Node Number (2)	Demand (cfs) (3)	Pressure (ft) (4)
Slack (cont.)	12	1.20	528.55
	13	-17.76	617.00
Fire at node 3	1	0.00	504.33
	2	4.00	423.99
	3	4.50	393.71
	4	3.43	386.86
	5	2.97	390.02
	6	2.40	387.50
	7	4.00	395.09
	8	0.00	417.14
	9	1.60	413.12
	10	1.60	430.55
	11	0.00	441.64
	12	1.60	439.80
	13	-26.10	617.00
Fire at node 8	1	0.00	487.14
	2	4.00	404.21
	3	2.08	379.55
	4	3.43	356.92
	5	2.97	356.54
	6	2.40	356.96
	7	4.00	363.03
	8	4.50	368.71
	9	1.60	365.55
	10	1.60	391.96
	11	0.00	404.36
	12	1.60	402.52
	13	-28.18	617.00

coefficient was computed based on 500 regressions for parameter C's. A sigma bound equal to 3 standard deviations was used as the first conditioning criteria to generate

error terms. The regressed C's were also conditioned to be between 65 and 145 (Hazen-William coefficients). This may introduce a small bias in the parameter estimates but it provides a means of introducing physical reasoning to accept the randomly generated error term. Table 6.3 shows estimates of the C's and their standard deviations for selected pipes. Appendix B presents results for all pipes.

Pipes 8 and 15 are the links with highest standard deviations in their C's. They carry low flow and are considered weak pipes. Standard deviations for these pipes were in the range of 20 units.

Pipe 1 is the most important link in the network since it is the only pipe connected to the source of supply. Its variance was small since an error in its estimate affected the pressure head estimates for all nodes. As the CV of the Q's increased from 1% to 25% it caused the standard deviation of C for pipe 1 to increase from 0.2 to 4.6, as can be seen in column (3) on Table 6.3. For the same variation in the CV of the Q's, the increase in the standard deviation of C for pipe 10 was from 2.7 ft to 9.7. The standard deviation increased by a factor of 3.6. It also can be seen from column (5) that the estimates of the C's for the same pipe worsen as the CV of the Q's increased.

The standard deviation of C for pipe 12 increased by the same factor observed for pipe 10, although there was more uncertainty associated with the estimates for this pipe (standard deviation equal to 4.0 ft for CV of Q's equal 1%).

TABLE 6.3

ESTIMATES OF PIPE ROUGHNESS C FOR DIFFERENT LEVELS OF
UNCERTAINTIES IN Q'S AND STD OF H'S EQUAL 0.25 FEET

CV Q (%) (1)	Statistics (2)	Roughness C Pipe #						Trace cov(C) (9)
		1 (3)	8 (4)	10 (5)	11 (6)	12 (7)	15 (8)	
1	mean C	109.7	108.8	110.0	109.7	108.8	84.1	1255
	std. C	0.2	20.7	2.7	1.5	4.0	17.9	
5	mean C	109.7	105.4	109.9	109.5	109.4	85.8	1681
	std. C	0.9	22.7	4.4	2.8	6.7	17.6	
10	mean C	109.6	108.1	109.2	109.2	109.7	92.2	2303
	std. C	1.8	23.1	6.0	4.0	9.4	21.2	
25	mean C	109.6	104.3	108.4	108.4	111.4	101.2	3846
	std. C	4.6	22.4	9.7	7.1	14.7	23.3	
TRUE C		110	110	110	110	110	110	

Uncertainties in Nodal Pressures

The next set of runs considered CV on Q equal 5% and a standard deviation of H's varying from 0.1 to 1.0 ft. Results are presented in Table 6.4. As the CV of Q or the standard deviation of H increased so did the standard deviation of the C's. The standard deviation of C for pipes 8 and 15 were high since they did not carry significant flow. These pipes are considered weak pipes in the system and they are less important than links carrying larger

TABLE 6.4

ESTIMATES OF PIPE ROUGHNESS C FOR DIFFERENT LEVELS OF
UNCERTAINTIES IN H'S AND CV OF Q'S EQUAL 5%

Std H (ft) (1)	Statistics (2)	Roughness C Pipe #						Trace cov(C) (9)
		1 (3)	8 (4)	10 (5)	11 (6)	12 (7)	15 (8)	
0.1	Mean C	109.9	107.9	110.0	109.9	109.9	91.8	
	Std. C	0.9	21.3	3.8	2.5	5.6	19.5	1459
0.25	Mean C	109.7	105.4	109.9	109.5	109.4	85.8	
	Std. C	0.9	22.7	4.4	2.8	6.7	17.6	1681
0.5	Mean C	109.0	106.9	108.5	107.6	109.2	83.9	
	Std. C	1.2	22.3	5.3	3.6	8.6	17.3	1992
1.0	Mean C	107.6	108.6	106.4	104.2	110.9	86.2	
	Std. C	1.5	23.2	6.1	4.6	11.7	19.3	2666
	TRUE C	110	110	110	110	110	110	

flows. Thus, a wider range of C values for pipes 8 and 15 do not greatly affect the pressure distribution throughout the system because the head drop through these links is not significant.

The full covariance matrix of estimated parameters and their correlations for the run with std of H's equal 0.25 ft (bases run), can be seen in Appendix B. Parameters C's were found to follow a normal distribution for all major pipes in the network as will be discussed later in this chapter.

Assessment of Calibration for a Different Demand

The demand selected to assess the calibration represents the flow requirements to fight a fire located at node 4. The magnitude of the flow at that node is 10 cfs. The remaining nodes in the network had their demands increased by 20% of their normal values. At this critical network operating condition we want to evaluate what the available nodal pressures will be. Table 6.5 shows the assigned flows for each node. The negative sign for the flow at node 13 indicates that it is a source of water supply for the network. Model prediction uncertainty was assessed by Monte Carlo and FOSM. Both methods considered correlation among parameter and the uncertainties associated with each parameter.

The trace of the covariance matrix of computed pressures was used as a measure of the calibration accuracy. Values from Table 6.6 show the assessment measure using parameters obtained from different calibration runs computed by the Monte Carlo method. Values presented are the average of 5 realizations. The technique for obtaining such values is explained in Chapter V. The probability of occurrence of the predicted pressures followed the normal distribution.

Each Monte Carlo realization represents results from 1000 network simulations. Mean pressures and corresponding standard deviations computed by FOSM are presented in Table

TABLE 6.5
ASSESSMENT LOAD FOR NETWORK 2

Condition (1)	Node Number (2)	Flow (cfs) (3)
Fire at node 4	1	0.000
	2	6.000
	3	3.120
	4	10.000
	5	4.452
	6	3.600
	7	6.000
	8	0.000
	9	2.400
	10	2.400
	11	0.000
	12	2.400
	13	-40.372

6.7 for all calibration runs. Table 6.8 summarizes the results obtained from previous runs and compares the assessment measure obtained by both methods. Appendix B includes results for all nodes.

Results from Monte Carlo and FOSM methods are very similar for most of the runs. The assessment measure computed by the two methods, as shown in columns (4) and (5) in Table 6.8, diverges more with the increase in the level of the input uncertainty. This fact is caused because the first order method does not consider higher order terms when approximating the variances of the nodal pressures.

TABLE 6.6
CALIBRATION ASSESSMENT FOR NETWORK 2
BY MONTE CARLO METHOD

Par. Estim.		Mean and Std of Pressures at Nodes (ft)							
CV (%) (1)	Q Std H (ft) (2)	Stat. (3)	Node Number					Trace Cov(Ha) (9)	
			1 (4)	4 (5)	7 (6)	9 (7)	11 (8)		
1	0.25	mean	492.9	202.0	238.4	279.6	346.5		
		std	1.1	2.5	2.1	2.0	1.9	50	
5	0.25	mean	492.8	201.7	238.1	279.4	346.5		
		std	3.8	8.5	7.8	7.0	6.0	598	
10	0.25	mean	492.6	201.6	237.2	279.2	346.4		
		std	7.5	16.5	15.5	13.9	11.8	2301	
25	0.25	mean	490.9	197.6	233.1	275.0	343.5		
		std	19.5	42.0	39.2	35.0	29.8	14834	
5	0.10	mean	493.5	203.7	239.2	280.9	347.8		
		std	3.8	8.4	7.7	6.9	5.9	578	
5	0.25	mean	492.8	201.7	238.1	279.4	346.5		
		std	3.8	8.5	7.8	7.0	6.0	598	
5	0.50	mean	490.0	196.1	233.5	274.5	341.6		
		std	5.0	10.1	9.5	9.3	8.6	974	
5	1.00	mean	483.9	184.5	222.4	262.9	330.1		
		std	6.9	13.0	12.4	12.8	12.8	1788	
True H's			494.3	206.5	239.9	282.1	348.9		

Therefore, the difference in the results computed by FOSM is accentuated with the increase of the nonlinearity effects introduced by the high uncertainty levels. This was a reason to incorporate a conditioning criteria to reject values from the tails of the distribution when generating random deviates. However, at high uncertainty levels, exact

TABLE 6.7
 CALIBRATION ASSESSMENT FOR NETWORK 2
 BY FIRST ORDER METHOD

Par. Estim.			Mean and Std of Pressures at Nodes (ft)						
CV (%) (1)	Q (ft) (2)	Std H (3)	Stat.	Node Number					Trace Cov(Ha) (9)
				1 (4)	4 (5)	7 (6)	9 (7)	11 (8)	
1	0.25	mean		492.9	202.0	238.6	279.6	346.5	
		std		1.1	2.6	2.2	2.0	1.8	53
5	0.25	mean		492.9	202.0	238.4	279.5	346.5	
		std		3.8	8.5	7.9	7.0	6.0	600
10	0.25	mean		492.8	202.1	237.7	279.5	346.5	
		std		7.7	17.0	16.0	14.3	12.2	2451
25	0.25	mean		492.5	202.8	236.7	278.6	346.0	
		std		19.7	43.2	40.6	36.3	30.9	15859
5	0.10	mean		493.6	204.0	239.5	281.0	347.9	
		std		3.8	8.6	7.9	7.0	6.0	608
5	0.25	mean		429.9	202.0	238.4	280.0	346.5	
		std		3.8	8.5	7.9	7.0	6.0	600
5	0.50	mean		490.0	196.4	233.6	274.6	341.6	
		std		5.0	10.2	9.6	9.3	8.6	982
5	1.00	mean		483.9	185.0	222.7	263.1	330.2	
		std		6.9	13.1	12.4	12.9	12.8	1801
True H's				494.3	206.5	239.9	282.1	348.9	

estimates are not of major concern if the measure is still consistent and robust. When the uncertainty is high, more data must be collected thus, as long as the measure shows the relative magnitude, it is adequate.

TABLE 6.8
COMPARISON OF CALIBRATION ASSESSMENT
COMPUTED BY DIFFERENT METHODS

Parameter Estimation			Calibration Assessment		
CV Q (%) (1)	Std H (ft) (2)	Trace cov(C) (3)	Trace cov(Ha) Monte Carlo (4)	(ft ²) FOSM (5)	
1	0.25	1255	50	53	
5	0.25	1681	598	600	
10	0.25	2303	2301	2451	
25	0.25	3846	14834	15859	
5	0.10	1459	578	608	
5	0.25	1681	598	600	
5	0.50	1992	974	982	
5	1.00	2666	1788	1801	

Impact of Different Assessment Loads

Different assessment loads can be selected according to the intended purpose of the calibration. This section shows how the measure of uncertainty in model results changes according to the loading condition used. Several assessment conditions will be studied using parameters estimated with CV of Q's equal to 5% and no uncertainty in the H's. The trace of the covariance matrix of the regressed C's in this case equals 1397 (Hazen-Williams roughness units).

The assessment loads considered fire fighting requirements at nodes 4, 7 and 10. Their magnitudes are presented in column (2) on Table 6.9. In all cases the demands at remaining nodes were increased by 20%. Column (3) shows the trace of the covariance matrix of computed pressures, $\text{trace cov}(H_a)$, which has been used as a measure of the calibration assessment. Other measures such as the average, minimum and maximum standard deviations of nodal pressures were also included in columns (4), (5), and (7) respectively. Columns (6) and (8) show the nodes having minimum and maximum standard deviations.

The first run in Table 6.9 contains a 10 cfs fire demand at node 4. The measure from column (3) shows a value of 541 ft^2 for the trace of the covariance matrix of the pressures. In the subsequent run, the magnitude of the fire was reduced to 8 cfs causing a decrease in the measure to 439 ft^2 . The third run has the same condition of the first, except that the fire was located at node 7 instead of node 4. This caused the measure to change from 541 ft^2 to 488 ft^2 . A fire with less magnitude, 5 cfs, was simulated in node 10 as listed in the fifth run. The measure of 427 ft^2 obtained in this run is closer to the 439 ft^2 obtained in run 2 when a larger flow was induced at node 4.

Results from Table 6.9 stress the importance of the locations and magnitude of high critical demands selected for the calibration assessment. As an example, when the assessment was performed for the first loading condition

TABLE 6.9
COMPARISON OF UNCERTAINTY MEASURES FOR
DIFFERENT ASSESSMENT LOADS

Assessment Load		Uncertainty in Predicted Pressures					
Critical Node	Demand (cfs)	Trace Cov(Ha) (ft ²)	Mean Std(Ha) (ft)	Minim. Std(Ha) (ft)	Node #	Maxim. Std(Ha) (ft)	Node #
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
4	10.0	541	6.7	3.7	1	8.1	6
4	8.0	439	6.0	3.4	1	7.1	6
7	10.0	488	6.4	3.5	1	7.5	6
7	8.0	398	5.8	3.2	1	6.5	10
10	5.0	427	5.9	3.3	1	6.6	10
10	3.8	377	5.6	3.1	1	6.2	6

from Table 6.2 (Normal demand) the trace of the covariance matrix of computed pressures was only 165 ft² compared to values in the range of 450 ft² observed for the assessment demands listed in Table 6.9.

Calibration with Measurements from Low Stress Conditions

This section investigates the effects of calibrating a network using measurements obtained by stressing the system to different levels. The first set of calibration runs used input data information from the five loading conditions described earlier and presented in Table 6.2. They

correspond to the runs under the title of normal stress demands in Table 6.10. The second set of runs considered as available input data the previous loads reduced to 60%. In these runs the nodal demands from Table 6.2 were multiplied by a factor of 0.6 and the corresponding nodal pressures were acquired for this new condition. This would correspond to take field measurements while inducing less flow at the fire and ordinary nodes. To reduce the stress in the ordinary nodes, the data collection efforts should be carried early in the morning when consumer demands are low.

Three calibration runs were performed under each condition described above. They correspond to different levels of uncertainty in the input data used to estimate parameters. These uncertainties are represented by the coefficient of variation of the nodal demands (CV of Q's) and by the standard deviation of the nodal pressures (std of H's). The same level of uncertainty was assigned for corresponding runs under both conditions of network stress. A constant CV of Q's and varying std of H's were considered for each of the three runs. The low stress condition had less uncertainty in the Q's since their standard deviations are scaled according to the magnitude of the mean value of the Q. Uncertainties for the H's, however, were at the same level for both stress conditions since their standard deviations are assigned independently of the mean value of the nodal pressure.

A CV of Q's equal to 5% was used to express the

TABLE 6.10

COMPARISON OF CALIBRATION PERFORMED WITH MEASUREMENTS
FROM DIFFERENT LEVELS OF NETWORK STRESS

Parameter Estimation			Calibration Assessment	
CV Q (%) (1)	Std H (ft) (2)	Trace cov(C) (3)	Trace cov (Ha) Load 1 (4)	(ft ²) Load 2 (5)
Normal Stress Demands				
5	0.10	1459	589	92
5	0.25	1681	614	91
5	0.50	1992	986	148
Low Stress Demands				
5	0.10	1239	551	83
5	0.25	2228	1515	228
5	0.50	2952	3262	492

uncertainty level of the Q's for all runs. The uncertainties in the measurement of the pressures correspond to standard deviation of the H's equal to 0.1 ft, 0.25 ft and 0.5 ft. The uncertainty level used in each run is listed in columns (1) and (2) on Table 6.10. Parameter uncertainty is shown in column (3).

Calibration assessment was performed for two assessment loads considering each calibration run. The first, Load 1, corresponds to the load previously utilized in this chapter and listed in Table 6.5. The second, Load 2, corresponds to the previous load with demands reduced to

60% of their original values. Calibration assessment for all runs were performed by FOSM method and are shown in columns (4) and (5) on Table 6.10.

Results from this investigation show that the calibration performed with low stressed demands gives better results at a small level of uncertainty in the measurements of the pressure heads. For example, considering the standard deviation of H's equal to 0.10 ft, the measure obtained with Load 1 was 551 ft², for the low stress demands, compared to 589 ft², for the normal stress demands. These values were 83 ft² and 92 ft², respectively, with Load 2. The standard deviation of H's equal to 0.10 ft was the only case in which the assessment measure was smaller when using calibrated parameters from low stress demands. The difference between the measures for different loads was in the range of 10%.

When the standard deviation of the H's was 0.25 ft, the measure obtained for Load 1 was 614 ft² for the calibration with normal stress demands compared to 1515 ft² for the less stress condition. The same measures considering Load 2 were 91 ft² and 228 ft². The difference was approximately 150% more for the measure obtained with the low stress demands. For the third run, with standard deviation of H's equal to 0.5 ft, the same difference increased to 230%. These results reveal the importance of stressing more the network demands when taken measurements to be used for calibration purposes. Since the pipes in the

network carry less flow the error in measurements become more significant relative to the system's losses when using low stress demands.

Different Amount of Available Information

This section investigates the quantity of the available information used to estimate parameters and its impact on parameter and prediction uncertainty. Two conditions were considered. The first compared uncertainties resulting from calibration performed using data from three and five loads. The second condition used information from five loads to estimate parameters but in each run a particular nodal demand was considered as unknown for all measured loads.

Parameter estimation runs in this section were completed after 500 regressions were performed to determine uncertainty in parameter C's. The uncertainty in the input data used to estimate parameters was the following: CV of Q's equal to 5% and standard deviation of H's equal to 0.25 ft. Table 6.11 show results from all conditions analyzed. Column (1) lists the number of loading conditions available to estimate parameters. Column (2) figures the node with the unknown demand. This demand in an extra variable to be identified by the estimation procedure.

The first run considered measurements under normal, peak and fire at node 8 conditions. The second and

TABLE 6.11
EFFECTS OF THE AMOUNT OF AVAILABLE INFORMATION
IN THE ASSESSMENT MEASURE

Assessment Measures						
Number of Loads (1)	Node w/ Missing Demand (2)	Trace cov(C) (3)	Trace cov(Ha) (ft ²) (4)	Average Std(Ha) (ft) (5)	Minim. Std(Ha) (ft) (6)	Maxim. Std(Ha) (ft) (7)
3	-	1896	950	8.9	4.8	10.7
5	-	1681	599	7.1	3.8	8.5
5	2	1878	35318	54.2	34.3	67.8
5	5	2077	5322	21.0	10.7	31.6
5	8	2137	29698	49.7	26.8	61.9
5	10	2205	3732	17.6	9.6	28.1

subsequent runs considered input data from all 5 loads as listed in Table 6.2. The number of equations that can be written for the regression matrix is 36 in the first case and 60 in the second case. The number of unknown parameters in both cases is equal to 16, corresponding to the number of pipes in the network.

The worth of the input data can be quantified by looking at the trace of the covariance matrix of the regressed parameters listed in column (3) on Table 6.11. It shows a value equal to 1681 obtained using 5 loads compared to 1896 with 3 loads. The uncertainty in the predictions also varied from 599 ft² to 950 ft² using 5 and 3 loads (See

column (4)). Other measures such as the average, minimum and maximum std (Ha) varied consistently with the changes in the amount of information as can be seen in columns (5) through (7).

The second set of runs shows the importance of the correct estimation of the demands for critical nodes. It is convenient to mention again that each node without demand data causes the loss of N_L equations in the regression matrix, where N_L is the number of loads. In this application the number of equations lost due to a missing nodal demand was 5. The number of available equations decreased from 60 to 55. The estimation performed with less information generates parameters with more uncertainty.

The trace of the covariance of C does not change much when information from different nodes is missing, as can be seen in column (3). However, the variances of estimates of C's for critical pipes may increase significantly and may influence the uncertainty in the resulting pressures. For example, if the demand at node 2 is not known (third run on Table 6.11), the measure in column (4) increases to 35318 ft² compared to 5322 ft² if the demand is not known at node 5. This happens because node 2 is located near the source of supply and at the intersection of three important links. Conversely, node 5 is located at the end of the network and is formed by the junction of only two less important pipes. This fact stresses the importance of the information at critical network locations and its contribution to the

magnitude of the measure of uncertainty in model predictions.

Network Parameterization

The effect of considering a group of pipes having the same C values was investigated. The purpose of parameterizing a network is to reduce the number of unknowns and consequently to increase the level of system overdetermination which improves the certainty of the regressed parameters.

The uncertainty presented in the input data for the runs in this section were a CV of Q's equal to 5% and standard deviation of H's equal to 1.0 ft. A total of 500 regressions were performed to estimate parameters for each run. The estimation procedure was conditioned to accept regressed C's ranging from 65 to 145 units. The first run was considered as a base run and it was completed in 72 hours using an IBM compatible personal-computer with a 486-25 MHz CPU.

Pipe 15 is a weak pipe in the network and has high uncertainty associated with the estimates of its roughness factor. In the subsequent runs this pipe was lumped with other pipes in an attempt to improve its estimates of C. Table 6.12 shows results for the base condition without any parameterization and three conditions with lumped parameters.

The second and third runs considered pipe 15 lumped

TABLE 6.12
EFFECTS OF NETWORK PARAMETERIZATION IN THE
ESTIMATES OF ROUGHNESS FACTOR FOR PIPES

Lump C at pipes (1)	Statistics (2)	Estimates of Roughness C Pipe #					Comput. Trace cov(C) (8)	Time (hr) (9)
		5 (3)	12 (4)	13 (5)	14 (6)	15 (7)		
-	mean C	89.4	110.9	105.8	109.8	86.2		
	std. C	19.0	11.7	19.5	13.3	19.3	2666	72
14, 15	mean C	93.5	107.0	104.5	101.0	101.0		
	std. C	21.2	10.7	20.2	11.2	11.2	2333	2.5
13, 15	mean C	91.6	118.6	83.5	106.0	83.5		
	std. C	20.4	8.6	15.4	11.4	15.7	2000	2.5
5, 13, 15	mean C	79.7	120.6	79.7	106.7	79.7		
	std. C	12.4	7.2	12.4	10.5	12.4	1352	-0.5
	TRUE C	110	110	110	110	110		

with pipes 14 and 13 respectively. The last run considered pipes 5, 13 and 15 lumped together. Column (8) from Table 6.12 shows the measure of uncertainty in the estimates of C. Column (9) shows the computational time to complete 500 valid regressions. It is apparent that the measure of uncertainty and the computational time decreases as more pipes are lumped to the same C.

When several pipes are considered as having the same roughness parameter, the number of unknowns in the system

reduces by the same number of pipes being parameterized while the number of equations remains the same. The improvement in the estimates is caused by the increase in the system's overdetermination level. Lumping a weak pipe with others having smaller variability contributes to reduce the variability of the first with the expense of increasing the variability of the other pipes.

Convergence of the Mean and Standard Deviation of C's

The number of Monte Carlo simulations performed to estimate the C's and their variability was based on the stochastic convergence of the mean and standard deviation of the C's observed for the base condition. This condition considered input data uncertainty equal to a CV of Q's at the 5% level and standard deviation for the H's equal to 0.25 ft, as listed in Tables 6.3 and 6.4. Most of the runs presented in this chapter were performed with 500 simulations. This number was enough for the convergence of the mean C and standard deviation C of the major pipes in the network. The convergence of these statistics for weak pipes would probably require a larger number of runs. However, the improvements obtained are not worth the effort since these links do not significantly contribute to improve the calibration assessment measure.

The program developed to estimate parameters has an option to compute the mean and standard deviations of the

C's after a specified number of simulations. These values can be saved and plotted to check the stochastic convergence of such variables.

Figures 6.2 and 6.3 show stochastic convergence of the mean and standard deviation of pipe 1 for the base run. It can be seen that convergence for both were obtained within less than 500 simulations. Figures 6.4 and 6.5 show the same analysis for pipe 10. This pipe has a higher degree of uncertainty associated with its roughness parameter and so it requires more simulations to converge. Stochastic convergence for other pipes considering the base run is included in Appendix E. Based on these results the use of 500 simulations was deemed as adequate.

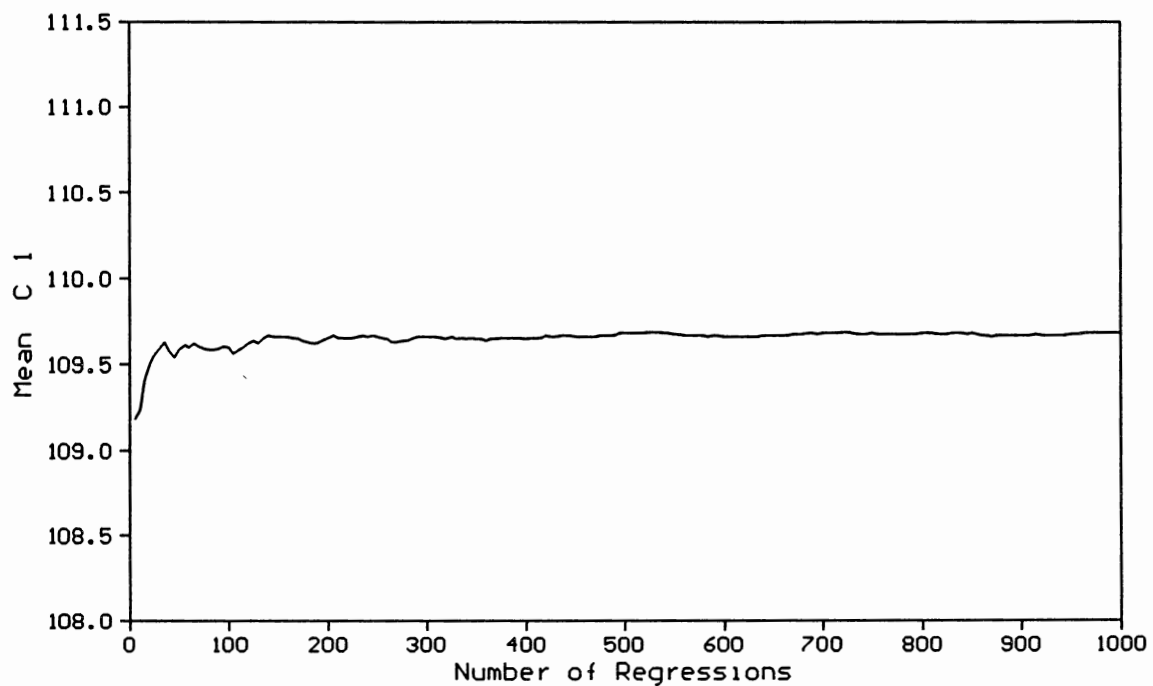


Figure 6.2. Stochastic Convergence of Mean C₁

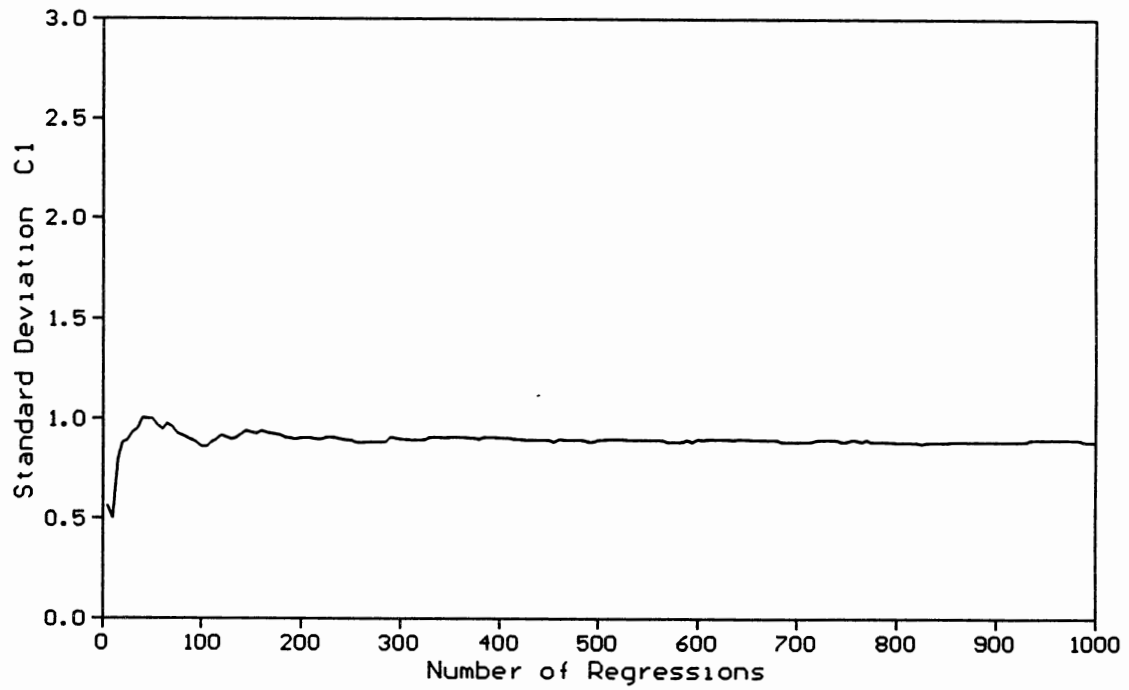


Figure 6.3. Stochastic Convergence of Std C_1

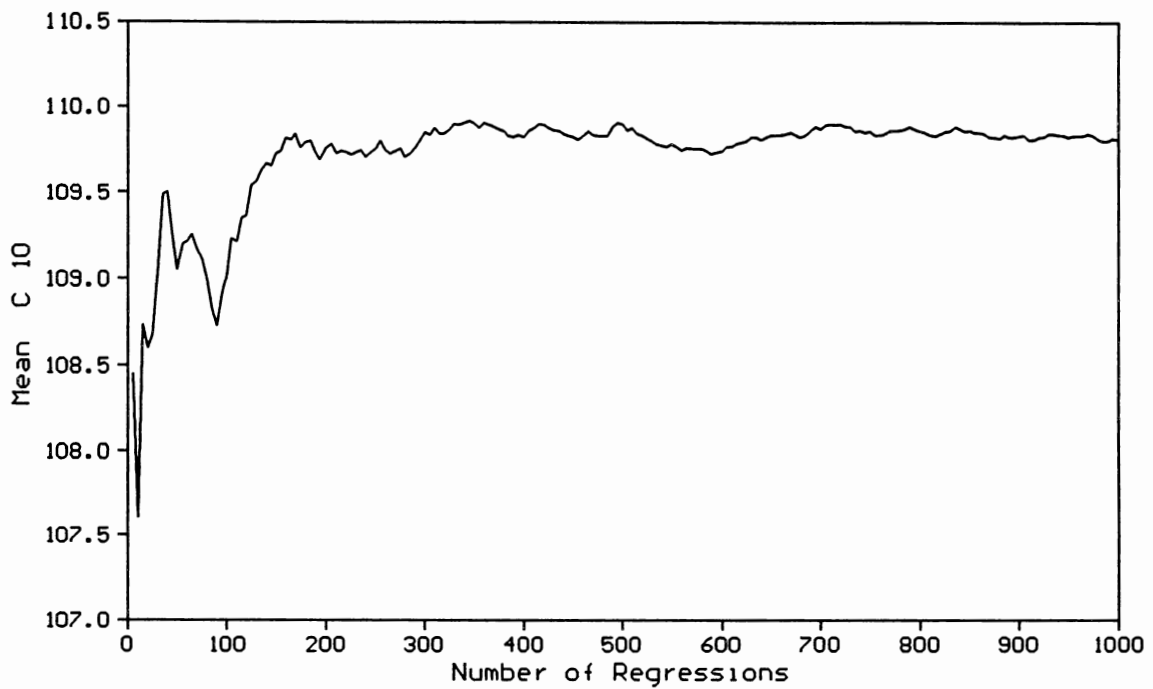


Figure 6.4. Stochastic Convergence of Mean C_{10}

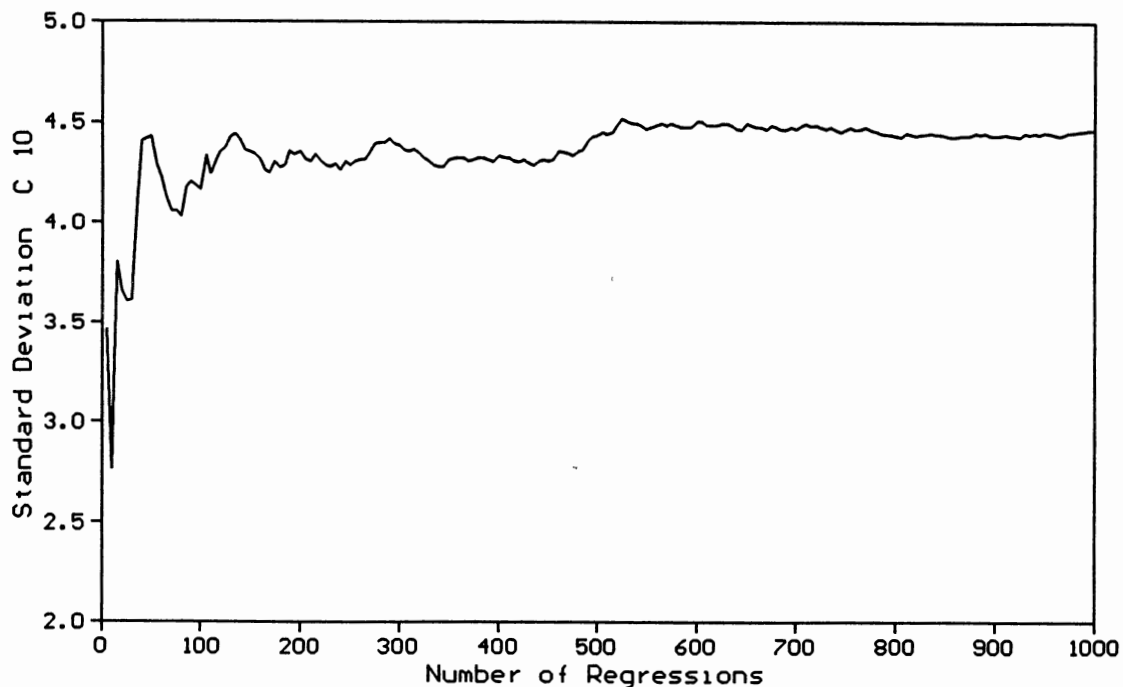


Figure 6.5. Stochastic Convergence of Std C_{10}

Approximation of the Distribution of Parameters

Parameter estimates from 500 regressions were used to approximate a distribution for the C's for the base run described in the previous section. The normal distribution was fitted to the data. Figures 6.6 and 6.7 show the distribution of C's from pipes 1 and 3, plotted against a normal distribution represented by the straight line. It can be seen that these data are well approximated by the normal distribution. A check of the normality for all pipes was performed in the same manner.

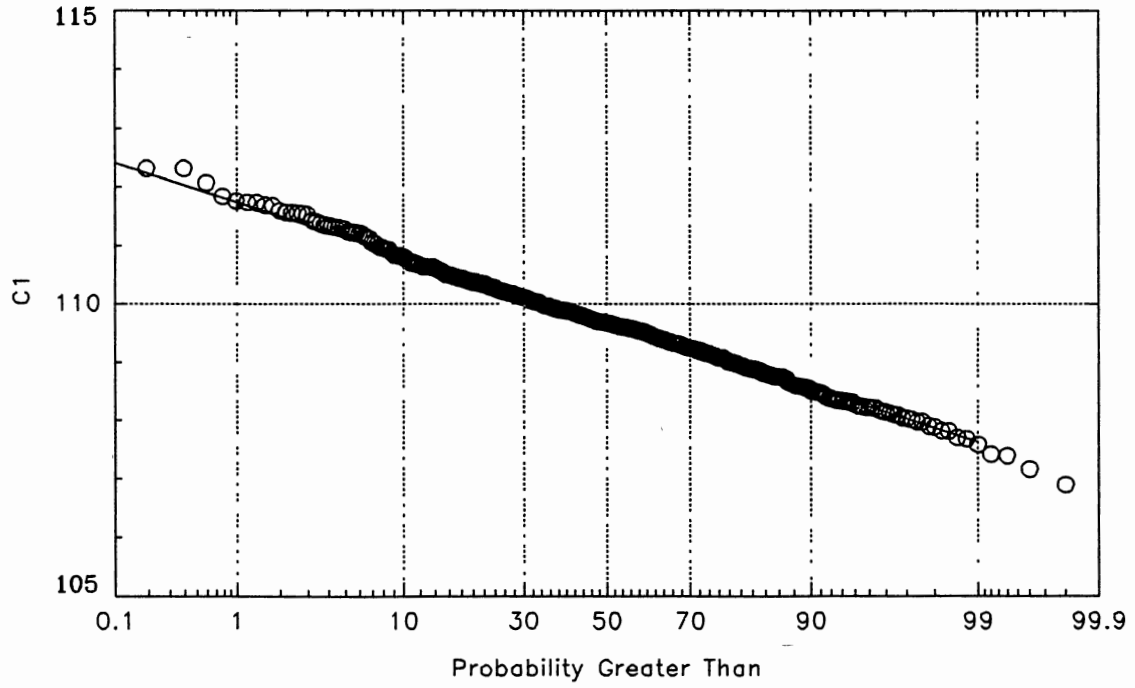


Figure 6.6. Probability Distribution of C_1

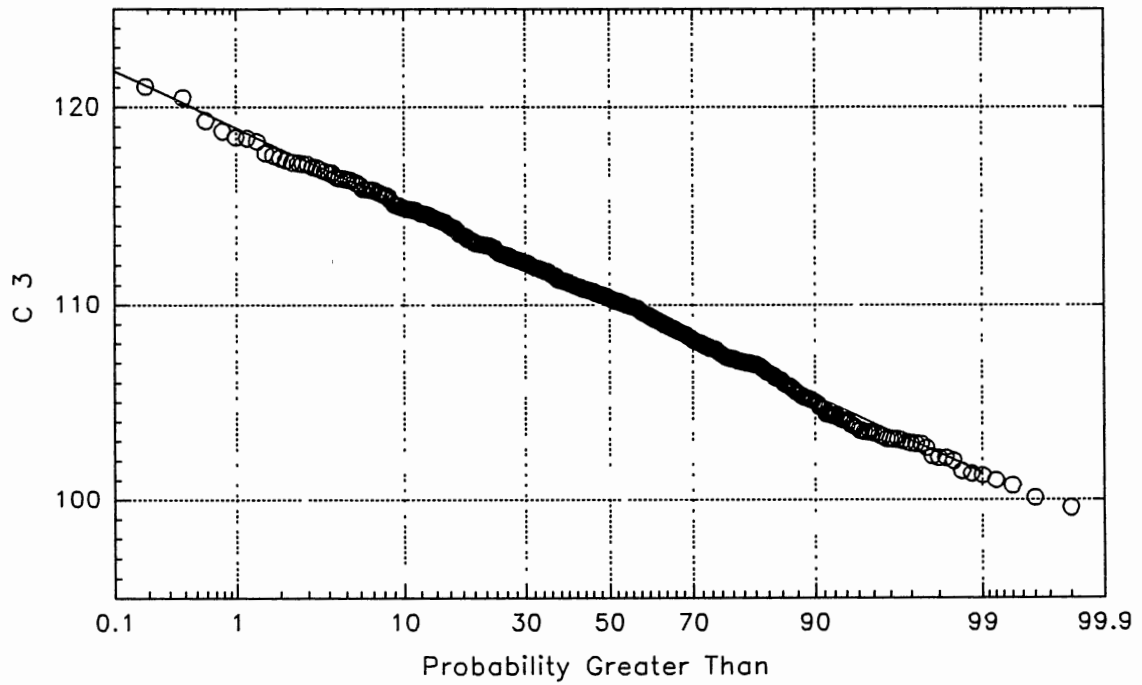


Figure 6.7. Probability Distribution of C_3

As an example, Figure 6.8 shows a fitting for pipe 9. Although the data deviated a little at the tails of the distribution, it did fit well at the center. It is noted that the estimate of pipe 9's C has higher standard deviation than those from pipes 1 and 3. Also, pipe 9 is located at the end of the network.

Estimates of C 's for important links more closely followed the normal distribution. Appendix F includes figures showing fittings for all pipes for the base condition. It was observed that only estimates for pipes 8 and 15 did not fit well by a normal distribution. Since

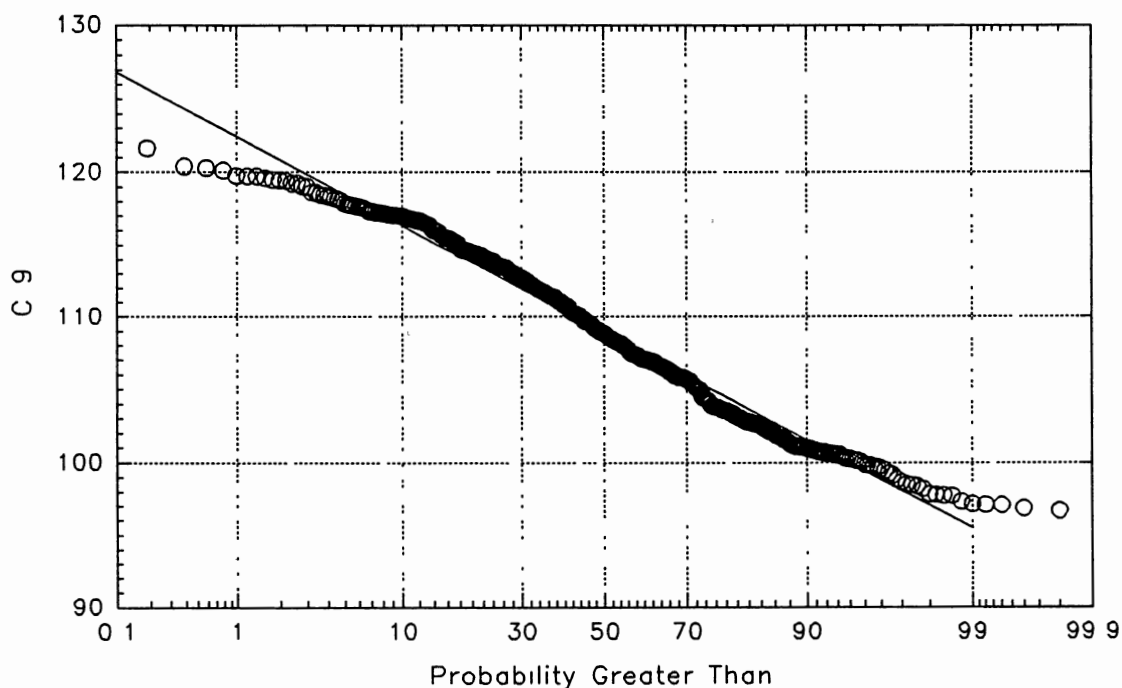


Figure 6.8. Probability Distribution of C_9

these pipes are weak in the network and carry low flow, their C's do not influence significantly the pressure distribution throughout the system. Therefore, the normal distribution was assumed to generate the error term for all pipe parameters when assessing the model predictions' uncertainty performed by the Monte Carlo method.

Distribution of Computed Pressures

Using data from each simulation performed for the calibration assessment by the Monte Carlo method, a normality check was conducted for the computed nodal pressure assuming that all C's were normally distributed. Figure 6.9 and 6.10 show plots of the pressure data from nodes 1 and 4 for the assessment load listed in Table 6.5, computed with parameters from the base run described earlier (See second run on Table 6.6). The data followed a straight line with the normal distribution.

Data from all nodes fit the normal distribution well as can be seen in the figures included in Appendix G. This fact is important since it allows confidence limits to be placed on the likelihood of the occurrence of a certain pressure at the nodes. The confidence limit is another way of expressing the uncertainty of the pressure in the node such that it can be understood by the layman.

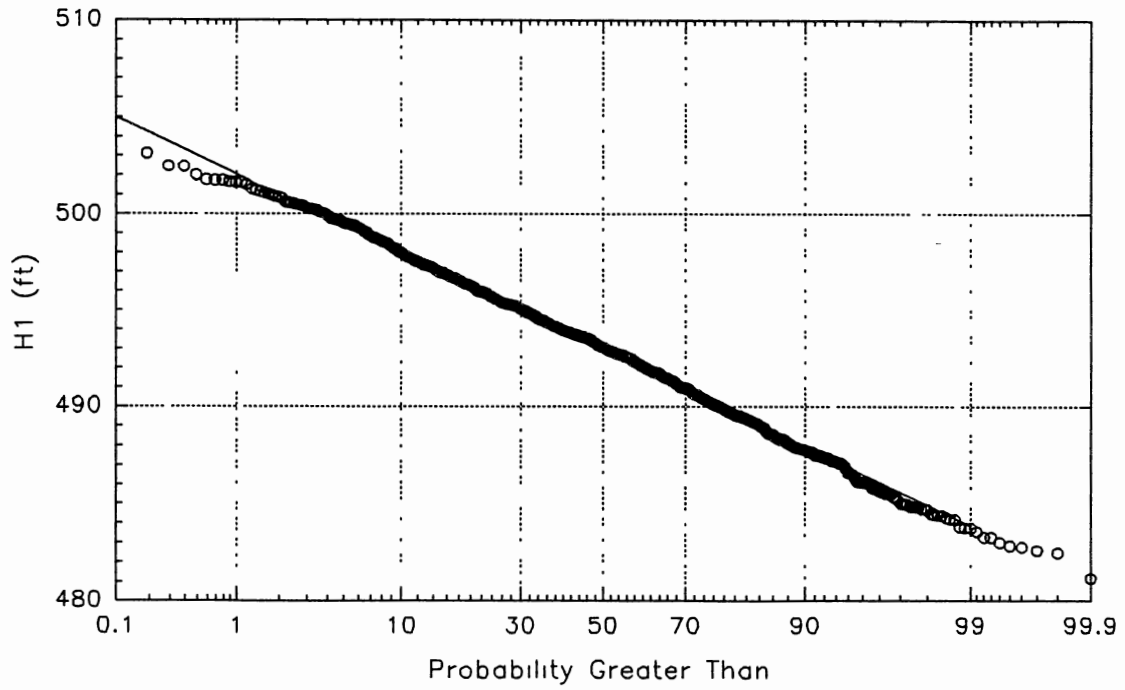


Figure 6.9. Probability Distribution of H_1

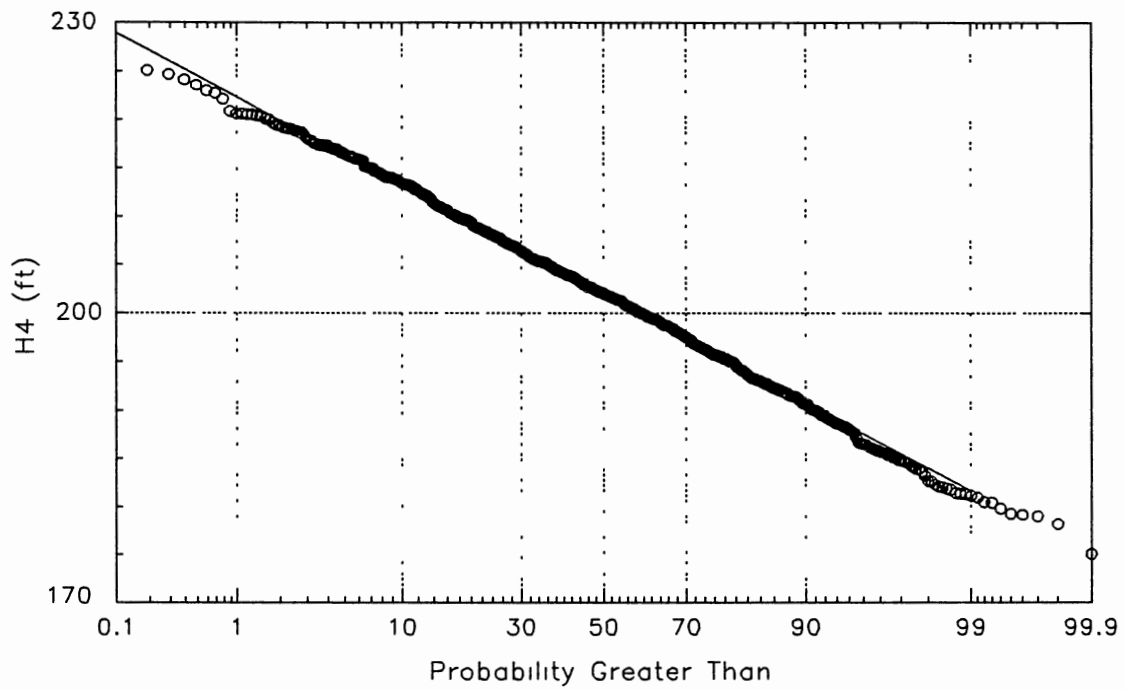


Figure 6.10. Probability Distribution of H_4

Appropriateness of the Assessment Measure

The variance of the computed nodal pressure was used to quantify the uncertainty in model predictions resulting from current knowledge of the calibrated parameters. The trace of the covariance matrix of the computed pressures, representing the sum of the variances of the pressures at all nodes in the network, is one way to express the total system variance. This quantity was selected as a measure to evaluate the calibration assessment. The purpose of this section is to justify the selection of the trace as a useful measure and to show its appropriateness to quantify uncertainties in model outputs for water distribution systems.

The first point considered to select the trace as a measure was the need of a single value to represent the output uncertainties. A single value is a fast and easy way of evaluating and reporting the improvements in the prediction uncertainty after each iteration is completed.

The applications developed in the present and previous chapters revealed the consistency of the measure represented by the trace. It changes in accordance with changes in the uncertainty levels of the input variables (Q's and H's). As the CV of Q or standard deviation of H increases, so does the trace. The magnitude of the changes in the trace is consistent with the changes in the uncertainty levels of the

measurements at critical nodes. These nodes are special by their strategic location in the network and because they carry high demands. Thus, the trace is an appropriate measure since it behaves well in regards to the type, magnitude and location of the changes in the uncertainty of input variables.

The second consideration for selecting the trace is that it can be computed by either Monte Carlo simulation or first order approximation. While the first is more accurate because it accounts for the system nonlinearity, the second method is very efficient in terms of computational time.

Results from both methods compared well as can be seen in Tables 5.6, 5.7 and 6.8. At high levels of uncertainty the difference in the trace computed by the two methods starts to deviate. However, the adopted measure can still be considered useful because, although not comparing exactly, the trace obtained by both methods are still close to each other. This makes this measure useful for the purposes of this study. If a node has high uncertainty associated with its predicted pressure, it is more important to identify that node and to work to reduce the uncertainty than to be able to compute the exact value of the uncertainty.

In summary, the trace was found to be a consistent and robust measure that can be used to quantify the model prediction uncertainties. As the amount, type and location of the uncertainties in the input variables changes, the

trace also changes in the appropriate direction. It also can be efficiently computed by FOSM and, if necessary, accurately determined by using Monte Carlo simulations. As a single measure, it represents the total system variance and can be used as a criteria to evaluate the improvements obtained in the computed pressures at each calibration iteration.

Other Useful Descriptive Criteria

Most of the applications developed in Chapters V and VI used the trace of the covariance matrix of computed pressure as the calibration assessment measure. This section discusses other measures that can be used as alternative criteria or to better represent different calibration objectives.

The standard deviation of the nodal pressures presented similar properties as those described above for the trace of the covariance matrix. By using individual nodal standard deviation as a measure, the modeler can assure to meet pressure prediction reliability at selected points in the network. Tables 6.9 and 6.11 present other useful measures as the average standard deviation and the minimum and maximum standard deviation of the nodal pressures. All of them are acceptable measures which could be used to meet particular calibration objectives.

Other measures could be defined to compare the uncertainty levels from different systems. This measure

would need to be normalized to reduce the effects of size and demands of the networks. Such measures could be defined as: (a) the average standard deviation divided by the total system headloss; (b) the average standard deviation divided by the total length of the pipe line; (c) the average standard deviation divided by the total system demand.

Another criteria to evaluate the calibration assessment could be defined based on the probability distribution of the computed pressure at the nodes. Since these pressures were well approximated by normal distributions, a criteria could be defined based on the confidence intervals to be placed on the mean of the predicted nodal pressure at a particular node.

Confidence intervals could be placed on critical nodes and used as criteria to determine if the current level of the calibration is accepted or if it is required to collect more field measurements to improve the calibration accuracy.

As an example, consider the predictions of the pressure at node 4 on Table 6.6. One may be interested in determining confidence intervals for the individual pressure considering the H's are normally distributed. Based on the properties of the normal distribution $(1-\alpha)\%$ of the pressures are within the limits

$$\bar{H} \pm z_{1-\frac{\alpha}{2}} S_H$$

where

$$\bar{H} \quad = \text{mean pressure (ft)}$$

$z_{1-\frac{\alpha}{2}}$ = comes from the standard normal distribution corresponding to the value of z such that the area under the standard normal density function from $-\infty$ to z is equal $1 - \alpha$
 S_H = standard deviation of the pressure (ft).

Considering the calibration run with CV of Q's equal 5% and standard deviation of H's equal 0.25 ft, the pressure at node 4 has a mean value of 201.8 ft and standard deviation of 8.5 ft (See column (5), Table 6.6). The 95% confidence interval for the predicted pressure at this node is computed as

$$\begin{aligned}
 H_U &= \bar{H} + z_{1-\frac{0.05}{2}} S_H \\
 &= 201.8 + 1.96 * 8.5 \\
 &= 218.5
 \end{aligned}$$

$$\begin{aligned}
 H_L &= \bar{H} - z_{1-\frac{0.05}{2}} S_H \\
 &= 201.8 - 1.96 * 8.5 \\
 &= 185.1
 \end{aligned}$$

The value of $Z_{0.975}$ equal 1.96 was obtained from the standard normal distribution. The lower and upper confidence limits computed as above are 185.1 ft and 218.5 ft. This means that the probability is 95% that the interval 185.1 ft and 218.5 ft contains the true mean of the pressure at node 4.

Considering the calibration run with CV of Q's equal 10% and std of H's equal 0.25 ft (See Table 6.6) the

pressure at node 4 has a mean value of 201.6 ft and std of 16.5 ft. The 95% confidence limits computed as above is 169.3 ft and 233.9 ft.

It can be seen that the 95% confidence interval increased from 33.4 ft in the first case to 64.6 ft as the uncertainty in the Q's used to estimate parameters increased from 5 to 10%.

CHAPTER VII

APPLICATION OF DATA COLLECTION PROCEDURE

This chapter presents applications of data collection procedures and the outer calibration loop presented in Figure 4.2. In the first section a system wide approach will be used to improve the estimates of the roughness parameters. In the second section, the improvements in the calibration assessment will be obtained by performing individual pipe tests.

System Wide Tests

This section presents an approach to identify ideal network conditions under which the new field measurements should be collected. The new measurements should be taken such that they will contribute to improve the current knowledge of the model parameters and to decrease the uncertainty in the results to be derived from the calibrated model.

For demonstration purposes, at the present stage, the calibration of the unknown roughness parameters has been performed using information from 4 measurement loads listed in Table 6.2: Normal, Peak, Fire at node 3, and fire at node 8. The external nodal demands from the above table were

estimated with a CV of Q equal to 5%. The corresponding pressure heads were measured or estimated with a standard deviation of 0.25 ft. With this information available, the estimates of parameter C were determined (See Table 7.1). These values reveal the current knowledge of the C for all pipes, represented by their mean and standard deviation. The trace of the covariance matrix of the estimates of C was found equal 1646.

The calibration exercise consists of improving the model's ability to predict the nodal pressures for the assessment load listed in Table 7.2. This condition reflects a fire with a magnitude of 8 cfs located at node 6. In addition to this fire load, the external demands at all other nodes are increased by 20% of their normal values which are listed in Table 6.2 as the first loading condition (Normal). The calibration assessment measure is represented by the trace of the covariance matrix of the model predicted pressures. This value and individual nodal standard deviations computed by FOSM method are presented in Table 7.3.

The objective of the calibration is to decrease the uncertainty level of the predicted pressure which corresponds in reducing the trace of the covariance matrix (717 ft²). In order to improve the assessment measure, more field data will be collected and used to augment the available information used to estimate parameters. The problem to be solved consists of providing guidance so that

TABLE 7.1
ESTIMATES OF PIPE ROUGHNESS C USING
INFORMATION FROM 4 MEASUREMENT LOADS

Pipe # (1)	Mean C (2)	Std C (3)
1	109.8	1.0
2	109.9	2.3
3	110.2	3.6
4	110.5	6.3
5	106.4	16.6
6	109.3	3.6
7	111.9	11.3
8	105.7	21.7
9	108.7	6.0
10	110.1	4.2
11	109.6	2.6
12	109.6	2.6
13	110.4	12.6
14	109.8	5.6
15	89.6	19.6
16	109.5	4.4
Trace Cov (C) = 1646		

the new collected data will add the most information. This guidance refers to the identification of network conditions (magnitude and location of the fire demands) to be induced in the real system when new measurements are taken.

TABLE 7.2
ASSESSMENT LOAD FOR DATA COLLECTION EXAMPLE

Condition (1)	Node Number (2)	Flow (cfs) (3)
Fire at node 6	1	0.000
	2	6.000
	3	3.120
	4	5.148
	5	4.452
	6	8.000
	7	6.000
	8	0.000
	9	2.400
	10	2.400
	11	0.000
	12	2.400
	13	-39.920

TABLE 7.3
CALIBRATION ASSESSMENT WITH PARAMETERS
ESTIMATED USING 4 MEASUREMENT LOADS

Node # (1)	Mean Ha (ft) (2)	Std Ha (ft) (3)
1	498.6	4.1
2	326.0	7.2
3	274.9	8.2
4	228.7	9.3
5	237.3	8.9
6	211.0	9.8
7	241.1	8.7
8	298.6	7.5
9	290.5	7.5
10	331.5	6.8
11	355.7	6.4
12	351.8	6.4

Trace Cov (Ha) = 717.0 ft²

Selection of Potential Measurement Loads

This step corresponds to blocks 4 and 5 in the diagram from Figure 4.2. The procedure is explained in detail in "Criteria to Select the Measurement Load" in Chapter IV.

In this example five potential network loadings have been determined as feasible conditions to be induced during the field experiments. They form the array Q_p . Each load corresponds to a vector Q_{p_i} which is composed of individual nodal demands $q_{p_{ij}}$, where j represents the node number. The magnitudes of these demands are listed in Table 7.4. The first load, Q_{p_1} corresponds to a fire flow equal 10 cfs at node 1 and the demands at other nodes reduced by 20% of their normal values. The second and third demands correspond to fire conditions of the same magnitude located at nodes 7 and 4 respectively. The above demands reflect single fire conditions located close to the source, in the middle, and at the extreme of the network. The fourth and fifth demands contain fire situations to be induced simultaneously at adjacent nodes in the east part of the network (nodes 4 and 5) and at the north and south parts of the network (nodes 2 and 9). The magnitude of the fire at node 9 is only 5 cfs because this node is a connection of pipes with small diameters.

As a first step the load Q_{p_1} is investigated. When the demands $q_{p_{1j}}$ are induced, the field tests will measure the

TABLE 7.4
POTENTIAL MEASUREMENT LOADS FOR DATA COLLECTION

Condition (1)	Node Number (2)	Predicted Demand (cfs) (3)	Estimated Pressure (ft) (4)
Q _{p1} Fire at Node 1	1	10.00	567.7
	2	4.00	500.7
	3	2.08	482.1
	4	3.43	467.8
	5	2.97	468.9
	6	2.40	467.8
	7	4.00	473.4
	8	0.00	490.4
	9	1.60	486.0
	10	1.60	500.1
	11	0.00	510.7
	12	1.60	508.8
	13	-33.68	747.0 (FGN)
Q _{p2} Fire at Node 7	1	0.00	603.6
	2	4.00	502.6
	3	2.08	476.3
	4	3.43	451.3
	5	2.97	456.5
	6	2.40	437.3
	7	10.00	437.7
	8	0.00	487.7
	9	1.60	484.5
	10	1.60	510.0
	11	0.00	522.3
	12	1.60	520.5
	13	-29.68	747.0 (FGN)
Q _{p3} Fire at Node 4	1	0.00	598.4
	2	4.00	494.3
	3	2.08	455.4
	4	10.00	411.2
	5	2.97	428.1
	6	2.40	429.6
	7	4.00	447.8
	8	0.00	477.4
	9	1.60	474.2
	10	1.60	501.4
	11	0.00	513.9

TABLE 7.4 (Continued)

Condition (1)	Node Number (2)	Predicted Demand (cfs) (3)	Estimated Pressure (ft) (4)
	12	1.60	512.0
	13	-30.25	747.0 (FGN)
Q_{p4}	1	0.00	549.4
Fires at	2	4.00	413.5
Nodes 4 and 5	3	2.08	357.3
	4	10.00	284.9
	5	8.00	284.9
	6	2.40	323.6
	7	4.00	350.3
	8	0.00	380.1
	9	1.60	378.0
	10	1.60	420.0
	11	0.00	434.3
	12	1.60	432.4
	13	-35.28	747.0 (FGN)
Q_{p5}	1	0.00	571.6
Fire at	2	10.00	444.2
Nodes 2 and 9	3	2.08	426.1
	4	3.43	412.6
	5	2.97	414.0
	6	2.40	412.6
	7	4.00	418.1
	8	0.00	436.7
	9	5.00	393.0
	10	1.60	459.6
	11	0.00	476.6
	12	1.60	474.7
	13	-33.08	747.0 (FGN)

corresponding nodal pressure heads. The new data on both demands and pressures will be added to augment the available

information which is used to estimate parameter C's.

The second step of the data collection module consists of predicting the worth of the information provided by the potential measurement load before conducting field measurements. Since the actual test was not yet conducted, the nodal pressures are not available and must be estimated. This task is accomplished by computing the pressures with KYPIPE using the current values of the C's from Table 7.1 and the demands from Q_{pi} . The estimated nodal pressures obtained by this process are listed in column (4) on Table 7.4 for the first iteration of the calibration loop. This value may change in subsequent iterations as the knowledge of the C's are improved.

The proposed measurement load Q_{pi} then supplements the available information Q_A and is used to estimate the unknown parameters. These projected parameters and their uncertainties are used to compute the assessment measure for the desired network condition.

The process is repeated for all loads from Q_p and a sensitivity vector is assembled (sensitivity vector A). This vector contains the difference in the assessment measure obtained with Q_A (trace cov (Ha) = 717 ft²) and the new measurement obtained when a potential load is considered to augment the available information (trace cov (Hp)). Table 7.5 lists results for the five potential measurement loads considered in this application. The trace of the cov (Hp) listed in column (5) was computed by FOSM. Column (6) shows

TABLE 7.5
SELECTION OF THE MEASUREMENT LOAD

Prediction of Uncertainty Measures					
Potential Load	Fire Node	Fire Demand (cfs)	Trace Cov(C)	Trace Cov(Hp) (ft ²)	Sensit. Vector A (ft ²)
(1)	(2)	(3)	(4)	(5)	(6)
Q _{p1}	1	10	1470	520	197
Q _{p2}	7	10	1271	632	85
Q _{p3}	4	10	1180	564	153
Q _{p4}	4, 5	10, 8	861	504	213
Q _{p5}	2, 9	10, 5	1163	553	164

the sensitivity vector A. It can be seen that adding the load with fires at nodes 4 and 5 (potential load Q_{p4}) causes the largest decrease in the assessment measure which reflects in the largest element of the sensitivity vector A; therefore, this load is selected for a more careful investigation regarding the individual nodal demands. Since this load has the largest total demand and the fire node is located at a distant point from the source, the largest head loss occurs when it occurs. The high losses would likely provide the most information to the analysis and affects critical links to the largest degree. Better estimates of roughness factor for critical links provide the largest decrease in the trace of the covariance of H_a.

The next step of the data collection process is to adjust individual nodal demands in the selected measurement load. For the purpose of this example we will restrict the investigation to the two nodes carrying fire demands which are nodes 4 and 5. After perturbing the individual demands, a new potential load is formed and the process is repeated. The potential load is incorporated into the available measurement; the parameters are estimated; and the assessment is performed. Table 7.6 presents results from this new step which considered both fire demands increased and reduced by 2 cfs. It can be seen that the load with fires at nodes 4 and 5 equal to 10 cfs provides the largest decrease in the assessment measure. This corresponds to the original fire demand at node 4 and the fire demand at node 5 increased by 2 cfs, as listed in column (3). A loading condition as close as possible to the one above selected (Q_{p4} , with adjusted flow at node 5) will be induced in the network at the time the new measurements are to be taken.

It is noted that the load with the largest total demand and causing the largest total head loss provided the most information. This result follows the standard practice, which suggests inducing large demands to cause the maximum head loss when collecting field data for calibration. This analysis also revealed a location which has not been previously considered.

TABLE 7.6
ADJUSTMENT OF NODAL DEMANDS FOR
THE SELECTED MEASUREMENT LOAD

Prediction of Uncertainty Measures			
Node 4 Demand (cfs) (1)	Node 5 Demand (cfs) (2)	Trace Cov(C) (3)	Trace Cov(Hp) (ft ²) (4)
10	8	861	504
12	8	1044	491
10	10	839	481
8	8	1009	490
10	6	1226	521

Collection of New Measurements

At this point the procedure would rely on the field data consisting of collecting field measurements of the actual induced flows in the network and the corresponding nodal pressures, Q_M and H_M . This corresponds to block 7 in the diagram on Figure 4.2. To generate this data the actual induced flows are assumed to be identical to those from the selected measurement load as listed in Table 7.7. The pressure heads are computed by KYPIPE using the true C's as shown in column (4) on Table 7.7.

The next step consists of estimating the parameters with the available information augmented by adding the new

TABLE 7.7
ADDITIONAL MEASUREMENT LOAD

Condition (1)	Node Number (2)	Demand (cfs) (3)	Pressure from KYPIPE (ft) (4)
Fire at	1	0.00	528.9
Nodes	2	4.00	379.2
4 and	3	2.08	316.7
5	4	10.00	233.5
	5	10.00	228.9
	6	2.40	272.6
	7	4.00	306.3
	8	0.00	340.0
	9	1.60	338.3
	10	1.60	386.5
	11	0.00	401.6
	12	1.60	399.8
	13	-37.28	747.0 (FGN)

measurements. New estimates of C's are presented in Table 7.8. By comparing these values with those from Table 7.1 a decrease in estimated parameter uncertainty is apparent. The improvements in the parameters will contribute to the decrease in the uncertainty of the predicted pressures as can be seen by comparing results from Tables 7.9 and 7.3. It is expected that at each new iteration the C's converge to their true values and their uncertainty reduces.

TABLE 7.8
ESTIMATE OF PIPE ROUGHNESS C USING
INFORMATION FROM 5 MEASUREMENT LOADS

Pipe # (1)	Mean C (2)	Std C (3)
1	109.8	0.8
2	109.9	2.1
3	110.2	3.2
4	110.5	5.1
5	107.0	10.2
6	110.0	2.1
7	109.9	9.3
8	109.4	16.4
9	109.7	5.1
10	109.8	3.5
11	109.7	2.5
12	109.5	6.3
13	109.5	12.2
14	108.9	5.4
15	106.7	9.8
16	109.2	4.1
Trace Cov (C) = 882		

At some point in the process, the assessment measure will satisfy the modeler's objective and the calibration process is complete. In a second case, none of the potential loads from Q_p can contribute to improvements in the assessment measure. If this happens, either improved data accuracy must be achieved, individual pipes can be examined, or the model is used as it is with its known predictions' uncertainty.

TABLE 7.9
 CALIBRATION ASSESSMENT WITH PARAMETERS
 ESTIMATED USING 5 MEASUREMENT LOADS

Node # (1)	Mean Ha (ft) (2)	Std Ha (ft) (3)
1	498.7	3.4
2	326.2	5.9
3	274.5	6.6
4	227.4	7.3
5	236.7	8.9
6	213.1	7.2
7	242.3	7.7
8	298.7	7.0
9	290.6	6.2
10	331.7	5.7
11	355.9	5.4
12	351.9	5.4
Trace Cov (Ha) = 472.2 ft ²		

Throughout the data collection process estimated C's are used to determine improvements in the assessment measure. The sensitivity vectors then are the estimates of the measurements true decrease and are likely conservative. As additional data is collected both the new value of C and Ha and their variances should improve. The above analysis, only accounts for reduction of the variance. A better estimate of the mean will further reduce the uncertainty in the results of this application. The data collection

procedure's goal is to identify useful field conditions for new measurement. The sensitivity vectors provide a comparison, which although not exact, does serve that purpose.

Individual Pipe Tests

This section investigates the effects of the improvements in the estimates of the roughness factor of individual pipes in the decrease of the measure of predictions' uncertainty. The network used to conduct experiments for this analysis is presented in Figure 6.1 and Table 6.1. The assessment load used is listed in Table 6.5. The calibration objective is to reduce the total system variance expressed by the trace of the covariance matrix of computed pressures, $\text{trace cov}(H_a)$.

The procedure starts by assigning a value for the estimates of the C and their uncertainty. This situation occurs at the beginning of the calibration efforts when no field measurements are yet available. At this point, the modeler seeks guidance to design data collection strategies and to plan the field experiments.

A value of C equal to 110 with a coefficient of variation of 10% was considered as the best guess for all 16 pipes from network 2. This corresponds to assigning a standard deviation of 11 units to all pipes. With this knowledge of the C 's, the measure of the total system variance for the predicted pressures was 32334 ft^2 as can be

seen in row 1, column (4) on Table 7.10. The average standard deviation for the pressures was 51.9 ft.

The selection of a pipe to conduct individual tests to improve the predicted pressures is based on the changes in the uncertainty measure which is used to assemble Gradient 2 described in chapter IV. First, individual pipe roughness improvements to a CV equal to 2% will be considered and the corresponding decrease in the trace cov (Ha) examined. Second, the pipe which is giving the largest contribution to decreasing the trace cov (Ha), corresponding to the link with the largest gradient, will be selected. Third, the contribution provided by the improvements in the selected pipe will be investigated.

Upon defining a level of uncertainty acceptable for the selected pipe, a step wise procedure follows. The improved pipe is held to its desired level and the process is repeated to identify the next pipe to be chosen for a carefully investigation.

Table 7.10 presents results used to select the best among pipes 1, 6, 11, 12, and 15. These pipes are located scattered throughout the system. Column (4) shows the measure obtained when improving the CV of the particular pipe to 2% while holding the uncertainty for remaining pipes at the 10% level. It can be seen that when the C for pipe 1 was improved the measure decreased to 7104 ft² compared to 29976 ft² observed for the same improvement in pipe 11. The magnitude of the improvements are not unexpected, since pipe

TABLE 7.10

COMPARISON OF CALIBRATION ASSESSMENT MEASURES FOR
DIFFERENT LEVELS OF PARAMETER UNCERTAINTY

Parameter Uncertainty			Calibration Assessment Measure			
Individual Pipe #	Pipe CV (%)	Other Pipes CV (%)	Trace Cov(Ha)	Average Std(Ha)	Minim. Std(Ha)	Maxim. Std(Ha)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
-	-	10	32334	51.9	46.8	53.6
1	2	10	7104	24.3	9.4	27.7
11	2	10	29976	50.0	46.8	53.3
12	2	10	32214	51.8	46.8	53.6
6	2	10	32250	51.8	46.8	53.6
15	2	10	32333	51.9	46.8	53.6
1	5	10	12623	32.4	23.4	35.0
1	1	10	6316	22.9	4.7	26.5
Cv C = 2% pipe 1						
11	2	10	4739	19.9	9.4	27.0
12	2	10	6984	24.1	9.4	27.7
6	2	10	7020	24.1	9.4	27.7
15	2	10	7103	24.3	9.4	27.7

1 carries the entire flow to the system. Pipes 6 and 15 have small head loss because they carry low flow and so, they have a minor impact on the pressure distribution throughout the system. Pipe 1 was, then, selected to be more carefully studied.

The effects of improvements in the uncertainty of the predictions measure can be assessed by performing the

analysis with different values for the CV of C for pipe 1. When the CV of C_1 was changed to 5% the measure changed to 12623 ft². This value reduced to 6316 ft² when the CV of C_1 was fixed as 1%. Based on the previous results, it was decided that a CV of C_1 equal 2% would be adequate corresponding to a standard deviation of 2.2 for the estimate of C_1 .

The next set of runs considered as starting point a CV of C_1 equal 2% and CV of C equal 10% for the remaining pipes. It can be seen that the next pipe to be select is pipe 11 since it contributes to the next largest decrease in the measure. The improvement in terms of the standard deviation of individual pipes (H_{a_i}) may not be significant, although, the decrease in the trace is high. Both terms should be examined and one may decide not to pursue these localized tests.

Gradient 2, reflecting the changes in the measure due to improvements in the uncertainty of individual pipe's roughness, is used to identify potential pipes to be studied. It corresponds to a point estimate at the current mean value of the C's. It may be that the means are not correct, but the relatively magnitude of these terms among pipes would be similar. Therefore, the analysis and decision process can continue as described here.

CHAPTER VIII

CONCLUSION AND FUTURE WORK

The increasing complexity of problems faced by hydraulic engineers necessitates obtaining optimized and fast solutions to hydraulic problems through the use of simulation models. It is of crucial importance that the model used is calibrated to represent the real system as close as possible. If models are not properly calibrated, their results may be biased and they may not reproduce the responses of the real system. It is a waste of valuable resources to base decisions on results that do not reflect reality of the physical system.

This research has addressed the problem of improving the calibration procedure for water distribution networks. It provides a framework to assist modelers in obtaining more confident decisions, a task that can be achieved when the calibration effort produces a model which best approximates the real system.

To date, only deterministic approaches have been used to calibrate distribution network simulation models. However, the modeling assumptions and the input information used to estimate parameters are not known with certainty. They contain errors which are propagated throughout the

steps performed for the calibration and to the results to be derived from the model. Quantification and a better understanding of the propagation of these errors is an important consideration in any modeling application.

Conclusion

A statistically based calibration methodology for a water distribution network model has been developed. The procedure considers three components of the modeling process: data collection, parameter estimation and calibration assessment. The process accounts for the uncertainty in measurements, their impact on model parameters and the effect of these uncertainties on the outputs of the network simulator. It also provides assistance in defining data collection strategies to improve the model predictive ability.

The roughness parameters of the pipes are estimated based on the uncertainty of the measurements of nodal pressures and estimation of their demands. Parameter uncertainty is transferred to model prediction uncertainty. A measure of the calibration accuracy is defined based on the trace of the covariance matrix of the computed nodal pressures. This measure is assessed by two methods: Monte Carlo simulations and first order second moment (FOSM) approximation of the variance. The calibration methodology was applied to two distribution networks under several loading conditions.

Most water distribution calibration procedures have been deterministically based and do not consider the uncertainties associated with modeling parameters and input variables. This research is an initial effort to quantify these uncertainties and to consider data collection as an integrated component which contributes to improve model predictive ability.

The major conclusions of this research are that: (a) a framework has been established to incorporate the error term in all steps of the modeling process of a water distribution system; (b) the trace of the covariance matrix of the computed pressures can be used as a consistent measure to represent in a single quantity, the uncertainties in model predictions; and (c) the FOSM approach can be used as an efficient method to compute prediction uncertainty for water distribution network models. Several practical results were also found including: (1) the level of the uncertainty of major pipes in the network affects significantly the measure of the calibration accuracy; (2) the need to stress the network to high demands to obtain more robust parameter estimates; (3) a linearity was observed in transferring uncertainties from parameter to model results in the two networks studied; (4) the sensitivity vectors and the gradients described in this dissertation can be used to guide future data collection efforts.

Future Work

The analysis developed showed major points concerning the importance of the input information used to calibrate the network model. It revealed some areas that require further research. These included:

1. Extend the parameter estimation procedure developed to use information from extended periods of time and to consider multiple sources of supply and storage.
2. Extend the procedure to handle (a) other representations for the roughness parameter, besides the Hazen-Williams factor and (b) the case when all heads are not known. These situations would require the use of non-linear regression or other interpolation schemes. In these cases a new source of errors are introduced in the process and they must be quantified.
3. Investigate the effects of the uncertainty introduced by common modeling practices such as network skeletonization (consider a simplified system) and network parameterization (reduce the number of parameters).
4. Investigate the effects of different error distributions in input variables on the distribution of the results of the network simulation model. Also, investigate the effect of correlation among parameters in the output uncertainties.
5. Apply sparse matrix techniques to avoid the propagation of round-off errors in the inversion of the regression

matrix used to estimate model parameters.

6. Investigate analytic approaches to approximate the variance of roughness parameters, such as the first order approximation.

REFERENCES

- American Society of Civil Engineering, *Pipeline Design for Water and Wastewater*, Report of the Task Committee on Engineering Practice in the Design of Pipelines, New York, 1975.
- Ang, A.H-S. and Tang, W.H., *Probability Concepts in Engineering Planning and Design: Volume 2*, John Wiley and Sons, New York, 1984.
- Bargiela, A., "An Algorithm for Observability Determination in Water-System State Estimation," IEE Proc., Part D, 1985, pp. 245-249.
- Bargiela, A. and Hainsworth, G.D., "Pressure and Flow Uncertainty in Water Systems," Journal of Water Resources Planning and Management Division, ASCE, Vol. 115, n. 2, March, 1989, pp. 212-229.
- Beck, J.V. and Arnold, K.J., *Parameter Estimation in Engineering and Science*, John Wiley and Sons, New York, 1977.
- Benjamin, J.R. and Cornell, C.A., *Probability, Statistics, and Decision for Civil Engineers*, McGraw-Hill, New York, 1970.
- Bhave, P.R., "Calibrating Water Distribution Network Models," Journal of Environmental Engineering Division, ASCE, Vol. 114, n. 1, Feb., 1988, pp. 120-136.
- Borah, A. and Haan, C.T., "Impact of Random Data Errors on Parameter Estimation in Hydrologic Modeling," Transactions American Society Agricultural Engineers, Vol. 34, n. 3, 1991, pp. 857-864.
- Boulos, P. and Ormsbee, L., "A Comprehensive Algorithm for Network Calibration," in Water Resources Planning and Management and Urban Water Resources, Proceeding of the 18th Annual Conference of the ASCE Water Resources Planning and Management Division, New Orleans, LA, May, 1991, pp. 949-954.
- Chavent, G., "Identifiability of Parameters in the Output Least Squares Formulation," in *Identifiability of Parametric Models*, edited by E. Walter, Pergamon Press,

New York, 1987.

Cleveland, T.G. and Yeh, W.W-G., "Sampling Network Design for Transport Parameter Identification," *Journal of Water Resources Planning and Management Division, ASCE*, Vol. 116, n. 6, Nov/Dec, 1990, pp. 764-783.

Cleveland, T.G. and Yeh, W.W-G., "Optimal Configuration and Scheduling of Ground-Water Tracer Test," *Journal of Water Resources Planning and Management Division, ASCE*, Vol. 117, n. 1, Jan/Feb, 1991, pp. 37-51.

Cohen, G. and Carpentier, P., "Leak Detention and Parameter Estimation," in *Computer Applications in Water Supply*, Vol. 1, edited by Colbeck and Orr, John Wiley & Sons, 1988.

Federov, V.V., *Theory of Optimal Experiments*, Translated and Edited by W.J. Studden and E.M. Klimko, Academic Press, New York, 1972.

Haan, C.T., *Statistical Methods in Hydrology*, The Iowa State University Press, Ames, Iowa, 1977.

Hamberg, D. and Shamir, U., "Schematic Models for Distribution System Design. I: Combination Concept," *Journal of Water Resources Planning and Management Division, ASCE*, Vol. 114, n. 2, March, 1988, pp. 129-140.

Hsu, N-S. and Yeh, W.W-G., "Optimum Experimental Design for Parameter Identification in Groundwater Hydrology," *Water Resource Research, AGU*, Vol. 25, n. 5, 1989, pp. 1025-1040.

Lansey, K.E., Duan, N., Mays, L. and Tung, Y-K., "Water Distribution Design Under Uncertainties," *Journal of Water Resources Planning and Management Division, ASCE*, Vol. 115, n. 5, Sep., 1989, pp. 630-645.

Lansey, K.E. and Basnet, C., "Parameter Estimation for Water Distribution Networks," *Journal of Water Resources Planning and Management Division, ASCE*, Vol. 117, n. 1, Jan./Feb., 1991, pp. 126-144.

Leroy, G.H., *Uncertainty Modeling of Water Distribution Systems*, unpublished, Ph.D. dissertation, University of Wisconsin-Madison, May, 1988.

Mays, L.W. and Tung, Y., *Hydrosystems Engineering & Management*, McGraw-Hill, New York, 1992.

- Nishikawa, T. and Yeh, W.W-G.; "Optimal Pumping Test Design for the Parameter Identification of Groundwater Systems," *Water Resources Research*, AGU, Vol. 25, n. 7, July, 1989, pp. 1737-1747.
- Ormsbee, L.E. and Wood, D.J., "Explicit Pipe Network Calibration," *Journal of Water Resources Planning and Management Division*, ASCE, Vol. 112, n. 2, April, 1986, pp. 166-182.
- Ormsbee, L.E., "Implicit Network Calibration," *Journal of Water Resources Planning and Management Division*, ASCE, Vol. 115, n. 2, March, 1989, pp. 243-257.
- Powell, R.S., Irving, M.R., Sterling, M.J.H. and Usman, A., "A Comparison of Three Real Time State Estimation Methods for On-Line Monitoring of Water Distribution Systems," in *Computer Applications in Water Supply*, Vol. 1, edited by Colbeck and Orr, John Wiley & Sons, 1988.
- Press, W.H., Flannery, B.P., Tenkolsky, S.A. and Vetterling, W.T., *Numerical Recipes, the Art of Scientific Computing*, Cambridge Press, Cambridge, MA, 1986.
- Rahal, C.M., Sterling, M.J.H. and Coulbeck, B., "Parameter Tuning for Simulation Models of Water Distribution Networks," *Proc. Instn. Civ. Engrs.*, Part 2, Sept., 1980, pp. 751-762.
- Rahal, C.M. and Sterling, M.J.H., "Dynamic Parameter Tuning for Water Distribution Network Models in an Extended Period Simulation," *Proc. Instn. Civ. Engrs.*, Part 2, March, 1981, pp. 151-164.
- Shamir, U. and Howard, C.D.D., "Water Distribution Systems Analysis," *Journal of Hydraulics Division*, ASCE, Vol. 94, n. HY1, Jan., 1968, pp. 219-234.
- Shamir, U. and Howard, C.D.D., "Engineering Analysis of Water-Distribution Systems," *Water Technology/ Distribution*, AWWA, Sep., 1977, pp. 510-514.
- Simmonds, L.E. and Laverty, G., "Performance Specifications for Water Meters," *American Water Works Association, Distribution Systems Actions and Innovations*, 1980, pp. 188-192.
- St. John, R.C. and Draper, N.R., "D-Optimality for Regression Designs: A Review," *Technometrics*, Vol. 17, n. 1, Feb., 1975, pp. 15-23.

- Sterling, M.J. and Powell, R.S., "Real Time State Estimation and Control for Water Distribution Networks," unpublished paper, 1989.
- Sun, N-Z. and Yeh, W.W-G., "Coupled Inverse Problems in Groundwater Modeling: 2. Identifiability and Experimental Design," *Water Resources Research*, AGU, Vol. 26, n. 10, Oct., 1990, pp. 2527-2540.
- Troutman, B.M., "Errors and Parameter Estimation in Precipitation-Runoff Modeling: 1. Theory," *Water Resources Research*, AGU, 1985, pp. 1195-1213.
- Tullis, J.P., *Hydraulics of Pipelines*, John Wiley & Sons, New York, 1989.
- Usman, A., Powell, R.S. and Sterling, M.J.H., "Comparison of Colebrook-White and Hazen-Williams Flow Modes in Real-Time Water Network Simulation," in *Computer Application in Water Supply*, Vol. 2, edited by B. Coulbec and C. Orr, John Wiley & Sons, 1988, pp. 21-37.
- Walski, T.M., "Why Calibrate Water Distribution System Models ?," *Water/Engineering & Management*, Oct., 1983, pp. 27-28.
- Walski, T.M., "Technique for Calibrating Network Models," *Journal of Water Resources Planning and Management Division*, ASCE, Vol. 109, n. 4, Oct., 1983, pp. 361-372.
- Walski, T.M., "Case Study: Pipe Network Model Calibration Issues," *Journal of Water Resources Planning and Management Division*, ASCE, Vol. 112, n. 2, April, 1986.
- Walski, T.M., "Does Your Model Really Model Your Water Distribution System ?," *Public Works*, June, 1987, pp. 65-66.
- Walter, E., *Identifiability of Parametric Models*, Pergamon Press, New York, 1987.
- Wolfe, M.P. and Koelling, C.P., *Basic Engineering and Scientific Programs for the IBM PC*, R.J. Brady Company, 1983.
- Wood, D.J., "Computer Analysis of Flow in Pipe Networks Including Extended Period Simulations," *User's Manual*, University of Kentucky, Lexington, Kentucky, 1981.
- Wunderlich, W. and Giles, J., "Review of Pipe Network Analysis," Report N. WR2B-2-900-190, Tennessee Valley Authority, Norris, TN, Sep., 1986.

- Yan, J. and Haan, C.T., "Multi-Objective Parameter Estimation for Hydrologic Models - Multiple Objective Programming," Transactions American Society Agricultural Engineers, Vol. 34, n. 3, 1991, pp. 848-856.
- Yeh, H-D., "Theis' Solution by Nonlinear Least-Squares and Finite-Difference Newton's Method," Ground Water, Vol. 25, n. 6, Nov.-Dec., 1987, pp. 710-715.
- Yeh, W.W-G., and Yoon, "Aquifer Parameter Identification with Optimum Dimension in Parameterization," Water Resources Research, AGU, Vol. 17, n. 3, June, 1981, pp. 664-672.
- Yeh, W.W-G., "Review of Parameter Identification Procedures in Groundwater Hydrology: The Inverse Problem," Water Resources Research, AGU, 1986, pp. 95-108.

APPENDICES

APPENDIX A

ADDITIONAL RESULTS FOR CHAPTER V

TABLE A.1

ESTIMATED PIPE ROUGHNESS C FOR DIFFERENT
LEVELS OF UNCERTAINTIES IN Q'S
(Complement of Table 5.3)

CV of Q	Statistics	PIPE NUMBER										
		1	2	3	4	5	6	7	8	9	10	11
1%	Mean C	99.9	100.0	100.0	99.9	99.9	99.9	99.9	99.9	100.0	99.9	99.9
	Std. C	0.9	1.1	1.5	0.6	0.4	0.4	0.7	1.0	1.4	0.6	0.2
5%	Mean C	99.9	99.7	99.9	100.0	99.9	99.9	99.9	99.8	100.2	99.8	99.9
	Std. C	4.3	5.1	7.7	3.6	2.0	1.9	3.3	5.2	7.1	3.0	1.1
10%	Mean C	100.6	99.1	101.3	99.0	99.5	100.3	99.8	99.8	100.0	99.7	99.9
	Std. C	9.1	11.5	15.6	6.8	4.1	3.6	6.2	10.0	13.2	5.8	2.2

TABLE A.2
 CALIBRATION ASSESSMENT FOR NETWORK 1 BY MONTE CARLO METHOD
 (Complement of Chapter V)

Calibration Condition		N O D E #							
		1	2	3	4	5	6	7	
		CV Q(%)							
Table 5.6 Monte Carlo	1	Mean H	189.9	145.8	180.8	185.5	182.2	191.9	183.7
		Std. H	0.04	0.40	0.80	0.05	0.08	0.03	0.08
	5	Mean H	189.9	145.6	180.8	187.3	182.2	191.9	185.7
		Std. H	0.02	2.03	0.42	0.428	0.39	0.16	0.394
	10	Mean H	189.9	145.2	180.8	187.3	182.2	191.9	183.6
		Std. H	0.40	4.00	0.82	0.497	0.748	0.32	0.794
		Std H(ft)							
Table 5.7 FOSM	1	Mean H	189.9	145.8	180.8	187.3	182.2	191.9	183.7
		Std. H	0.04	0.41	0.08	0.05	0.08	0.03	0.08
	5	Mean H	189.9	145.73	180.8	187.3	182.2	191.9	183.7
		Std. H	0.21	2.1	0.44	0.26	0.4	0.17	0.41
	10	Mean H	189.9	145.53	180.8	187.3	182.2	191.9	183.7
		Std. H	0.42	4.13	0.85	0.52	0.78	0.34	0.83

APPENDIX B

ADDITIONAL RESULTS FOR CHAPTER VI

TABLE B.1

ESTIMATES OF PIPES ROUGHNESS C FOR DIFFERENT
LEVELS OF UNCERTAINTY IN CALIBRATION DATA
(Complement of Tables 6.3 and 6.4)

Data in	Calibration Condition	PIPE #															# Runs				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Accepted	Total		
Std H = 0.25 ft																					
CV Q(%)																					
Table 6.3	1	Mean C	109.7	109.71	110.3	110.7	106.4	109.0	110.2	108.8	109.5	110.0	109.7	108.8	111.6	109.5	84.11	108.5	500	8963	
		Std. C	0.25	1.16	2.94	5.65	14.54	2.71	9.61	20.72	5.05	2.71	1.54	3.99	7.66	5.33	17.91	3.75			
		5	Mean C	109.7	109.86	110.2	110.4	105.7	109.1	111.5	105.4	108.9	109.9	109.5	109.4	110.4	109.7	85.77	108.8	500	8515
			Std. C	0.89	2.27	3.77	6.70	17.59	3.86	11.65	22.66	5.78	4.43	2.79	6.67	13.24	5.95	17.61	4.12		
	10	Mean C	109.6	110.0	110.4	111.0	103.8	108.6	110.6	108.1	109.7	109.2	109.2	109.7	109.3	109.5	92.2	108.8	500	11119	
		Std. C	1.79	3.64	5.14	7.90	20.09	5.71	13.26	23.07	7.01	6.03	3.97	9.35	17.89	7.26	21.16	6.64			
	25	Mean C	109.6	110.6	110.4	110.6	106.0	110.0	109.5	104.3	108.4	108.4	108.4	111.4	104.8	109.5	101.2	108.5	500	76473	
		Std. C	4.59	7.25	10.24	14.49	22.07	11.88	19.06	22.41	12.43	9.72	7.13	14.73	22.39	14.30	23.27	14.15			
CV Q = 5%																					
Std H(ft)																					
Table 6.4	0.1	Mean C	109.9	109.9	109.9	109.9	109.2	109.6	111.0	107.9	109.4	110.0	109.8	109.9	109.9	110.1	91.8	109.8	500	1353	
		Std. C	0.90	2.06	2.60	4.12	15.06	3.36	11.05	21.27	5.65	3.76	2.49	5.56	11.42	3.67	19.51	3.20			
		0.25	Mean C	109.7	109.9	110.2	110.4	105.7	109.1	111.5	105.4	108.9	109.9	109.5	109.4	110.4	109.7	85.8	108.8	500	8515
			Std. C	0.89	2.27	3.77	6.70	17.59	3.86	11.65	22.66	5.78	4.43	2.79	6.67	13.24	5.95	17.61	4.12		
	0.5	Mean C	109.0	110.2	111.2	112.8	94.7	107.0	109.9	106.9	108.6	108.5	107.6	109.2	109.9	108.8	83.9	104.1	500	60796	
		Std. C	1.15	2.64	4.92	9.07	18.56	4.70	11.74	22.30	6.14	5.32	3.63	8.61	16.55	9.10	17.31	8.14			
	1	Mean C	107.6	110.4	110.3	111.7	89.4	105.8	107.4	108.6	108.1	106.4	104.2	110.9	105.8	109.8	86.2	93.2	500	1786480	
		Std. C	1.53	3.28	6.93	13.14	18.97	5.43	12.38	23.19	6.33	6.09	4.64	11.74	19.53	13.32	19.34	13.49			

TABLE B.2
 COVARIANCE MATRIX OF ESTIMATED C'S FOR CV OF
 Q EQUAL 5% AND STD OF H EQUAL 0.25 FT

C1	ROUGHNESS PARAMETER OF PIPES							
	C2	C3	C4	C5	C6	C7	C8	
C1	.78781E+00							
C2	.74303E+00	.51311E+01						
C3	.83364E+00	.25171E+01	.14232E+02					
C4	.93120E+00	.36871E+01	.23537E+02	.44888E+02				
C5	.57501E+00	-.61250E+01	-.33910E+02	-.61286E+02	.30927E+03			
C6	.98167E+00	-.69338E+00	-.69014E+01	-.13214E+02	.50845E+02	.14900E+02		
C7	.90301E+00	.75554E+00	.19986E+01	.35481E+01	-.65342E+01	-.25594E+01	.13568E+03	
C8	-.87696E+00	-.57253E+00	-.50877E+01	-.78325E+01	.13674E+02	.41905E+01	-.24231E+03	.51343E+03
C9	.20620E+00	-.10970E+00	-.95047E+00	-.13085E+01	.29511E+01	.94898E+00	-.57930E+02	.12143E+03
C10	.10549E+01	-.70763E+01	-.15603E+01	-.31935E+01	.12565E+02	.41443E+01	.18430E+02	-.37802E+02
C11	.82768E+00	-.47201E+01	-.12453E+01	-.24138E+01	.85789E+01	.29774E+01	.11905E+01	-.14718E+01
C12	.57879E+00	.11199E+02	-.87387E+01	-.16385E+02	.22670E+02	.60171E+01	.24211E+00	.39781E+01
C13	.11806E+01	-.22732E+02	.47044E+01	.88214E+01	-.39536E+02	-.95202E+01	-.61004E+01	.36803E+01
C14	.61995E+00	-.40831E+00	-.12756E+02	-.24268E+02	.26765E+01	.21279E+01	-.21610E+01	.54444E+01
C15	.81399E-01	.35759E+01	-.95861E+01	-.18804E+02	-.23634E+02	-.57355E+01	-.87215E+01	.17902E+02
C16	.27546E+00	-.69247E-01	.39170E+00	.21664E+00	-.10578E+01	.25521E+00	.65256E+00	-.31043E+01

TABLE B.2 (Continued)

C9	ROUGHNESS PARAMETER OF PIPES							
	C10	C11	C12	C13	C14	C15	C16	
C9	.33459E+02							
C10	-.89431E+01	.19614E+02						
C11	.45132E+00	.11223E+02	.77828E+01					
C12	.76279E+00	-.19582E+02	-.12819E+02	.44431E+02				
C13	.82481E+00	.46936E+02	.31209E+02	-.74527E+02	.17538E+03			
C14	.93142E+00	.24402E+01	.18835E+01	.13513E+02	.61212E+01	.35351E+02		
C15	.32750E+01	-.77963E+01	-.42375E+01	.21346E+02	-.10289E+02	.30186E+02	.30996E+03	
C16	-.70389E-01	.34420E+00	.53754E+00	-.66530E+00	-.11517E+01	-.19384E+01	.50052E+00	.16940E+02

TABLE B.3
CORRELATION AMONG PARAMETERS FOR CV OF Q
EQUAL 5% AND STD OF H EQUAL 0.25 FT

C1	ROUGHNESS PARAMETER OF PIPES							
	C2	C3	C4	C5	C6	C7	C8	
C1	.10000E+01							
C2	.36957E+00	.10000E+01						
C3	.24896E+00	.29455E+00	.10000E+01					
C4	.15659E+00	.24295E+00	.93122E+00	.10000E+01				
C5	.36838E-01	-.15376E+00	-.51113E+00	-.52015E+00	.10000E+01			
C6	.28653E+00	-.79302E-01	-.47393E+00	-.51097E+00	.74902E+00	.10000E+01		
C7	.87342E-01	.28635E-01	.45482E-01	.45464E-01	-.31899E-01	-.56924E-01	.10000E+01	
C8	-.43604E-01	-.11155E-01	-.59517E-01	-.51594E-01	.34316E-01	.47911E-01	-.91807E+00	.10000E+01
C9	.40163E-01	-.83725E-02	-.43556E-01	-.33763E-01	.29011E-01	.42503E-01	-.85978E+00	.92647E+00
C10	.26837E+00	-.70537E+00	-.93388E-01	-.10763E+00	.16133E+00	.24243E+00	.35726E+00	-.37670E+00
C11	.33426E+00	-.74692E+00	-.11833E+00	-.12914E+00	.17486E+00	.27649E+00	.36634E-01	-.23284E-01
C12	.97829E-01	.74168E+00	-.34751E+00	-.36690E+00	.19340E+00	.23386E+00	.31182E-02	.26339E-01
C13	.10044E+00	-.75778E+00	.94162E-01	.99422E-01	-.16976E+00	-.18624E+00	-.39547E-01	.12264E-01
C14	.11747E+00	-.30316E-01	-.56871E+00	-.60921E+00	.25598E-01	.92719E-01	-.31203E-01	.40411E-01
C15	.52090E-02	.89667E-01	-.14433E+00	-.15941E+00	-.76334E-01	-.84399E-01	-.42529E-01	.44876E-01
C16	.75404E-01	-.74274E-02	.25227E-01	.78562E-02	-.14615E-01	.16064E-01	.13611E-01	-.33286E-01

TABLE B.3 (Continued)

C9	ROUGHNESS PARAMETER OF PIPES							
	C10	C11	C12	C13	C14	C15	C16	
C9	.10000E+01							
C10	-.34910E+00	.10000E+01						
C11	.27968E-01	.90835E+00	.10000E+01					
C12	.19784E-01	-.66334E+00	-.68935E+00	.10000E+01				
C13	.10767E-01	.80026E+00	.84473E+00	-.84427E+00	.10000E+01			
C14	.27082E-01	.92670E-01	.11355E+00	.34095E+00	.77739E-01	.10000E+01		
C15	.32159E-01	-.99990E-01	-.86276E-01	.18190E+00	-.44129E-01	.28837E+00	.10000E+01	
C16	-.29566E-02	.18883E-01	.46815E-01	-.24250E-01	-.21129E-01	-.79209E-01	.69073E-02	.10000E+01

TABLE B.4

CALIBRATION ASSESSMENT FOR NETWORK 2 BY MONTE CARLO METHOD
(Complement of Table 6.7)

Calibration Condition		N O D E N U M B E R												TRACE	
		1	2	3	4	5	6	7	8	9	10	11	12		
Cv	Q														
%	Std H (ft)														
1	0.25	Mean H	492.9	316.5	258.4	202.0	216.2	214.4	238.4	287.6	279.6	322.1	346.5	342.5	50
		Std. H	1.1	1.7	2.0	2.5	2.5	2.5	2.1	2.0	2.0	1.9	1.8	1.9	
5	0.25	Mean H	492.8	316.4	258.3	201.7	216.1	214.0	238.1	287.5	279.4	322.1	346.5	342.5	599
		Std. H	3.8	6.6	7.5	8.5	8.4	8.4	7.9	6.9	7.0	6.4	6.0	6.0	
10	0.25	Mean H	492.6	315.8	257.7	201.6	216.1	212.7	237.2	287.3	279.2	322.0	346.4	343.2	2301
		Std. H	7.5	12.8	14.6	16.5	16.4	16.4	15.5	13.7	13.9	12.5	11.8	11.9	
25	0.25	Mean H	490.9	313.0	254.3	197.6	212.6	207.9	233.1	283.9	275.0	318.3	343.5	339.3	14835
		Std. H	19.0	32.9	37.3	42.0	41.8	41.5	39.2	34.6	35.0	31.9	29.8	30.0	
5	0.10	Mean H	493.6	317.4	259.7	203.7	217.6	215.1	239.2	288.9	280.9	233.4	347.8	343.8	578
		Std. H	3.8	6.5	7.5	8.4	8.1	8.2	7.7	6.8	6.9	6.3	5.9	5.9	
5	0.25	Mean H	492.8	316.4	258.3	201.7	216.1	214.0	238.1	287.5	279.4	322.1	346.5	342.5	599
		Std. H	3.8	6.6	7.5	8.5	8.4	8.4	7.9	6.9	7.0	6.4	6.0	6.0	
5	0.50	Mean H	490.0	312.1	253.1	196.1	211.7	209.1	233.5	282.8	274.5	317.0	341.6	337.2	974
		Std. H	5.0	7.9	9.0	10.1	10.3	10.2	9.5	9.2	9.3	8.9	8.6	8.8	
5	1.00	Mean H	485.9	303.1	243.1	184.5	200.3	197.6	222.4	271.4	262.9	305.1	330.1	324.6	1788
		Std. H	6.9	10.3	11.5	13.0	13.0	13.1	12.4	12.7	12.8	12.9	12.8	13.6	

TABLE B.5

CALIBRATION ASSESSMENT FOR NETWORK 2 BY FIRST ORDER METHOD
(Complement of Table 6.7)

Calibration Condition			1	2	3	4	N O D E N U M B E R								TRACE
Cv Q %	Std H (ft)						5	6	7	8	9	10	11	12	
1	0.25	Mean H	492.9	316.5	258.3	202.0	216.0	214.7	238.6	287.6	279.6	322.1	346.5	342.5	53
		Std. H	1.1	1.7	2.1	2.6	2.5	2.7	2.2	2.0	2.0	1.9	1.8	1.9	
5	0.25	Mean H	492.9	316.5	258.4	202.0	216.1	214.4	238.4	287.6	279.5	322.2	346.5	342.6	600
		Std. H	3.8	6.6	7.5	8.5	8.4	8.4	7.9	6.9	7.0	6.4	6.0	6.0	
10	0.25	Mean H	492.8	316.1	258.1	202.1	216.2	213.4	237.7	287.5	279.5	322.2	346.5	342.5	2451
		Std. H	7.7	13.2	15.0	17.0	17.0	17.0	16.0	14.1	14.3	12.9	12.2	12.3	
25	0.25	Mean H	492.5	315.9	258.4	202.8	216.7	212.3	236.7	286.9	278.6	321.5	346.0	341.9	15859
		Std. H	19.7	34.0	38.5	43.2	42.9	43.1	40.6	35.9	36.3	33.0	30.9	31.1	
5	0.10	Mean H	493.6	317.6	259.9	204.0	217.6	215.4	239.5	289.0	281.0	232.5	347.9	344.0	608
		Std. H	3.8	6.6	7.6	8.6	8.4	8.5	7.9	7.0	7.0	6.4	6.0	6.1	
5	0.25	Mean H	492.9	316.5	258.4	202.0	216.1	214.4	238.4	287.6	279.5	322.2	346.5	342.6	600
		Std. H	3.8	6.6	7.5	8.5	8.4	8.4	7.9	6.9	7.0	6.4	6.0	6.0	
5	0.50	Mean H	490.0	312.1	253.2	196.4	211.7	209.3	233.6	282.8	274.6	317.0	341.6	337.3	982
		Std. H	5.0	8.0	9.1	10.2	10.3	10.2	9.6	9.2	9.3	9.0	8.6	8.9	
5	1.00	Mean H	483.9	303.2	243.2	185.0	200.8	198.2	222.7	271.6	263.1	305.3	330.2	324.9	1801
		Std. H	6.9	10.3	11.5	13.1	13.1	13.1	12.4	12.7	12.9	13.0	12.8	13.7	

APPENDIX C

GENERATION OF MULTIVARIATE DEVIATES

GENERATION OF MULTIVARIATE NORMAL
RANDOM VARIABLES (Haan, 1977)

Multivariate normally distributed random variables can be generated so that the means, variances and correlations are preserved for all of the variables. The starting point is the correlation matrix \underline{R} of the random variables. The equation

$$(\underline{R} - \lambda \underline{I}) \underline{a} = 0$$

is solved to obtain the characteristic roots λ and vectors \underline{a} of \underline{R} . Another matrix \underline{Z} is generated as

$$\underline{Z} = (\underline{z}_1, \underline{z}_2, \dots, \underline{z}_n)$$

where \underline{z}_i is a vector composed of elements $z_{i,j}$. The elements $z_{i,j}$ are generated from a normal distribution with a mean of 0 and a variance of λ_j .

$$z_{i,j} \sim N(0, \lambda_j)$$

A matrix \underline{X} is then computed from

$$\underline{X} = \underline{Z} \underline{A}'$$

The elements of \underline{X} are $x_{i,j}$ and are normally distributed with a mean 0 and a variance of 1. Further

$$\underline{X}' \underline{X} = (n-1) \underline{R}$$

Finally $y_{i,j}$ is computed from

$$y_{i,j} = \sigma_j x_{i,j} + \mu_j$$

The $y_{i,j}$ are normally distributed with a mean of μ_j and a variance of σ_j^2 . The matrix \underline{Y} is made of elements $y_{i,j}$ and has

the property

$$\underline{Y}'\underline{Y} = (n-1)\underline{S}$$

where \underline{S} is the variance-covariance matrix that is to be preserved in the generation process. Thus the matrix \underline{Y} contains the desired simulated random variables that have the correct means, variances and correlations.

If multivariate lognormally distributed random variables are desired, the matrices \underline{R} and \underline{S} must refer to the logarithms of the variables. An additional step is required to get the antilog of the generated data. this step is to compute a matrix \underline{U} composed of elements $u_{i,j}$ determined from

$$u_{i,j} = \exp(y_{i,j})$$

APPENDIX D
PROGRAM LISTING

D1. List of programs and subroutines

Routine*	Objective
JVPIPE5	Main program for parameter estimation
JVASSMC	Main program for calibration assessment by Monte Carlo method
JVASSFO	Main program for calibration assessment by first order approximation
SPPC	Compute Physical Parameters for pipes
SMAKEB	Generate constant vector B(external nodal demands)
SNCON	Identify pipes connected to a node
SMAKEX	Generate coefficient matrix for regression
SSTAT	Compute mean and variance
SPERTQ	Introduce noise in Q's
SPERTH2	Introduce noise in H's
SPERTH3	Compute DC/DH
SMREGIA1	Perform multiple-linear regression
SCORR	Compute statistics of parameter estimates
SAVEC	Save estimates of C after each regression
SFLOAD	Compute component loadings
TRED2	Reduce a symmetric matrix to a tridiagonal form
TQLI	Find eigenvalues and eigenvectors
EIGSRT	Sort the eigenvalues
SPRTMAT	Prints a matrix or vector
DECOMPOS	Decompose a matrix into lower and upper form using scaled partial pivoting
SAUGMAT	Augment the coefficient matrix of regression to include intercept coefficient
SINVMAT	Inverts a matrix
SMULTMAT	Perform matrix multiplication
SOLVE	Solves a system of linear equation
STRANMAT	Compute the transpose of a matrix
SCONDR	Check criteria to accept estimates of C in the conditional Monte Carlo method
SNOISEC	Generates normally distributed noise and perturb C
GASDEV	Generates a gaussian deviate
STATJV	Compute correlation matrix
INDATA**	Reads geometry file (KYPIPE format)
GENMPNL**	Interface subroutine for KYPIPE
SIM**	Perform network hydraulic simulation

* Some routines are modifications of routines found in Press et al. (1986) and Wolfe and Koelling (1983).

** From Lansey and Basnet (1991).

D2. Program to Perform Parameter Estimation

```

C *****
C * PROGRAM : JVPIPE5 *
C * AUTHOR : JOSE VICENTE GRANATO DE ARAUJO *
C * DATE : APRIL 8, 1990 (JVPIPE) *
C * Oklahoma State University *
C * *
C * Last Revision Dec 29, 1991 *
C * *
C * REMARKS : This program calls subroutine Smregia to *
C * find the least squares estimates for coef. C *
C * by solving: *
C *  $q_i = c_1x_{i1} + c_2x_{i2} + \dots + c_px_{ip}$  *
C * *
C * VARIABLES *
C * ===== *
C * N : NUMBER OF OBSERVATIONS FOR REGRESSION (NN*NL) *
C * NP : NUMBER OF VARIABLES IN THE EQUATION *
C * NUMPERT : NUMBER OF PERTURBATIONS FOR Q *
C * ICONS 0 = REGRESSION THROUGH (0,0) *
C * 1 = COMPUTES Y INTERCEPT *
C * X(N,NP) : MATRIX OF COEFFICIENTS OF IND. VARIABLES *
C * Y(N) : VECTOR OF COEFF. OF DEPENDENT VARIABLE *
C * B(NP+ICONS) : COEFFICIENTS OF REGRESSION EQUATION *
C * LIMIT : ARRAY DIMENSION (must be > N) *
C * TOLER : DEFINE TOLERANCE TO APPROXIMATE ZERO *
C * NINP : # OF INPUT FILE *
C * NOUT : # OF OUTPUT FILE *
C * NOUC : # OF OUTPUT FILE TO SAVE C FOR STATISTICS *
C * NOUH : # OF OUTPUT FILE TO SAVE h perturbed " *
C * ISAVEC = 1 TO SAVE C PERTURBED IN THE FILE "SAVEC.DAT" *
C * ISAVEH = 1 SAVE PERTURBED H's FOR STATISTICS *
C * TITLE : TITLE OF RUN *
C * QJV(NN*NL) : CONSTANT VECTOR FOR REGRESSION(ORIGINAL) *
C * QP(NN*NL) : CONSTANT VECTOR FOR REGRESSION(PERTURBED) *
C * CJV : regressed C's *
C * CSAVE( , ) : SAVE VALUES OF Cjv's TO COMPUTE STATISTICS *
C * DIMENSIONS (NUMPERT,NP) *
C * CSTAT(2,NP) : CONTAINS STATISTICS(MEAN,VAR) OF C's *
C * CV : COEFF. OF VARIATION =  $S_x/X_{mean}$  to perturb  $Q_i$  *
C * IQH = 0 PERTURB Q *
C * = 1 PERTURB H *
C * = 2 PERTURB Q & H *
C * = 3 Compute dC/dHm *
C * = 4 Compute dC/dQ *
C * = 5 Compute dC/dM = [dC/dHm,dC/dQ] *
C * XNODE(NL,NN,2) ( , 1)= Q(nl,nn); ( , 2)= H(nl,nn) *
C * xmstd(i,j) : Std to perturb node j from load i *
C * STDI : Stand. Dev. used to pert H converted to psig *
C * STD : Stand. Dev. used to pert H in feet *
C * nstda : Label for type of statistics 1=mean, 2=std *
C * ndrop(i) : Vector with nodes do drop equation *
C * nundrop : Number of nodes to be drop (must be >0) *
C * xmcv(i,j) : Coef. of variation for node j, from load i *
C * indlq(i) : contains the load # corresponding to Q(i) *
C * indnq(i) : contains the node # corresponding to Q(i) *
C * icv : indicates the selected CV type to perturb Qs *
C * nsave : output file # to save results for later use *
C * fsave : output file name to save results *
C * igap : interval to compute mean & std to check conver *
C * iseedjv : seed to generate random deviates *
C * 0 to initialize from clock *
C * dcdh(np) : save column of dc/dh (or dC/dQ) matrix *
C * icondr : 1=condition regression, 0=not conditioned *

```



```

c      *      icondq : 1=condition noise on Q, 0=not conditioned      *
c      *      icondh : 1=condition noise on H, 0=not conditioned      *
c      *      xnstdq : # of std to condition noise on Q                *
c      *      qmax5 : Q limit to compute std of Q in spertq           *
c      *      xnstdh : # of std to condition noise on H                *
c      *      clowb,cupb : lower and upper bounds to accept C         *
c      *      nviocl,nviocu : # of violations for lower and upper bounds *
c      *      nvioql,nvioqu : # of violations of Q's                   *
c      *      nvioq,nvioh : # of violations of H's                     *
C      *      INPUT FILE                                              *
C      *      =====                                                *
C      *      TITLE (A60)                                              *
C      *      NL, NN, NP, INODE(NN,2), PIPE(NP,2), IPIPE(NP,2)*      *
C      *      XNODE(NL,NN,2)                                           *
C      *****
C      IMPLICIT REAL*8(A-H,O-Z)
C      INTEGER NUMP,NUMN
C      INTEGER JX,IP,M,JA,JB,IX,MPL,KC,KCLO,KIP,KPI,JIJ,JJJ

PARAMETER (LIMIT=70,NUMP=51,NUMN=51,NUML=5,NPERT=1000)
DIMENSION XNODE(NUML,NUMN,2),INODE(NUMN,8),IPIPE(NUMP,2),
*X(LIMIT,LIMIT),PIPE(NUMP,2),PIPEK(NUMP),QJV(LIMIT),CJV(LIMIT),
*QP(LIMIT),CSAVE(NPERT,20),CSTAT(2,20),XNODEP(NUML,NUMN,2),
*choice(6),nsta(3),ndrop(10),indlq(limit),indnq(limit),
*xmncv(numl,numn),dcdh(nump),XL(LIMIT,LIMIT),LVEC(nump),
*xmstd(numl,numn)
COMMON /jvga/NOUT,nouh,iseedjv

COMMON/SINDAT/JE(100),JG(100)
COMMON /SHAR/ BM(1100), IX(2200), IP(100,13), AL(100), GFH(700),
. AM(3800), JP(100), JX(4500)
COMMON/SSIMIN/ JTAN(50),EMAX(50),EMIN(50),DIAM(50),NFUL(50),
1 LY(32),LZ(32),E(100),KPI(100),JD(100),JF(100)
1 ,MPL(4500),NA(80),NB(80),JJUN(80),JP(80),
1 NPO(50),NJO(50),L10(10),L11(10),L12(10),XGRD(10),XGRD1(10),
1 XGRD2(10),DDQ,IOUT,JSKIP,KJ,KTEP,MAXT,NJOUT,NMOM,NXX
1 ,NPOUT,NPRP,NQ,NR,NTANK,NTEP,P,SW,TPER,TPPP
COMMON/SALL/QEXTK(50),AA(50),BB(50),CC(50),DD(50),EE(50),FF(50),
1 KIP(1000),JIJ(1000),KC(100),B(100),C(100),D(100),Q(100),JC(100)
1 ,R(100),S(100),V(100),ENGY(100),JJI(100),KCLO(100),JA(100),JB(100)
1 ,GG(50),A1,A2,A3,A4,CQ,I3,KN,NEPS,TIME,TNEXT,TPER1,TTT,UU,NJFIX
1 ,IEPS,KK,CHECK,ZQ(60),NABEL,NPUMP,NCODE
COMMON/PENAL/YMIN(100),YMAX(100)
COMMON/SNAME2/M(100)
COMMON /SINREV/ BI(100),JFIX(50),TNCD,JNCG,NPCG,NNP,LABEL,FAC,
1 XPER,NQEX
COMMON /INITBK/ INIT,LASTCL

CHARACTER*1 NFILEIN*20,NFILEOU*20,NFILEOC*20,TITLE*70,choice*5
character*11 TIMEOUT,DATEOUT*8,nfileoh*20,fsave*20,FSAVEM*20,
*FSAVEST*20,findata*20,fdcdh*20,fdcdq*20
data choice/'Q','H','Q&H','dC/dH','dC/dQ','dC/dM'/,
*nsta/'mean','std','var'/

c----These lines are from nlcode setup(Wadsop)
do 1 j=1,1100
  bm(j)=0.0
  ix(j)=0
1 continue
  init=1
  igap=5
  deltaq=0.0
  deltax=0.0
  nvioqu=0
  nviocl=0
  CV=0.0
  stdi=0.0
  xnstdq=0.0
  xnstdh=0.0
  nviohl=0
  nviohu=0
  nvioql=0
  nvioqu=0
  WRITE(*,8001)
8001 FORMAT(1H1,/,5X,' LEAST SQUARES ESTIMATION OF COEFFICIENTS FOR A'

```

```

*, ' PREDICTIVE EQUATION ', //, 5X, ' Author: Jose Vicente Araujo',
*5X, 'PROGRAM JVPIPE4'/)
WRITE(*,*) ' INPUT FILE NAME FOR OPTIONS AND FLOWS(or device)? '
READ (*,2)NFILEIN
2 FORMAT(A20)
write(*,*)' OUTPUT FILE NAME (or device)? '
READ (*,2)NFILEOU
4 WRITE(*,*) ' Noise on Q(0), H(1), Q&H(2), or compute gradients(3)
* ?'
READ(*,*)IQH
IF ((IQH .GT. 3) .OR. (IQH .LT. 0))GOTO 4
IF (IQH .EQ. 3)THEN
290 continue
write(*,*) ' Select gradient: dC/dH(3), dC/dQ(4) or dC/dM(5)?'
read(*,*)iqh
if( (iqh.lt.3) .or. (iqh.gt.5))goto 290
if((iqh .eq. 3) .or. (iqh .eq.5))then
291 continue
write(*,*)' Compute sensitivity matrix dC/dH '
write(*,*)' Give Delta H : '
read(*,*)deltah
if(deltah .eq. 0.0)goto 291
endif
if((iqh .eq. 4) .or. (iqh .eq.5))then
292 continue
write(*,*)' Compute sensitivity matrix dC/dQ '
write(*,*)' Give Delta Q : '
read(*,*)deltaq
if(deltaq .eq. 0.0)goto 292
endif
nundrop=1
ndrop(1)=nn
ENDIF
NINP=5
NOUT=6
NOUC=7
NOUH=8
nindata=9
nsave=66
NSAVEM=67
NSAVEST=68
ndcdh=61
ndcdq=62
fdcdh='for061.dat'
fdcdq='for062.dat'
findata='for009.dat'
fsave='for066.dat'
FSAVEM='FOR067.DAT'
FSAVEST='FOR068.DAT'
c-----Read Geometry data from Kypipe format
WRITE(*,*)' Reading Network Geometry from For009 '
open(unit=nindata,file=findata,status='old')
open(45)
open(unit=10,file='indata.out')
call indata
close(nindata)
close(45)
close(10)
c-----check if units=cfs
if (nq .ne. 0)then
write(*,*)' *****
write(*,*)' Units must be in CFS (nq=0) in file For009'
write(*,*)' *****
stop
endif
OPEN(UNIT=NINP,FILE=NFILEIN,STATUS='OLD')
OPEN(UNIT=NOUT,FILE=NFILEOU)
C*****
C READ INPUT VARIABLES
C*****
WRITE(*,*) ' ENTER TITLE OF THE PROBLEM(MAX60):'
3 FORMAT(A70)
READ (NINP,3)TITLE
WRITE(*,*) ' NUMBER OF LOADS (NL) : '
READ (NINP,*)NL

```

```

ccc      WRITE(*,*) ' NUMBER OF NODES (NN) : '
ccc      READ (NINP,*)NN
ccc      WRITE(*,*) ' NUMBER OF PIPES (NP) : '
ccc      READ (NINP,*)NP
c----Get values from indata
c----(This is valid when there is only one FGN present in the system)
      NN=KN+1
      NP=KK
      if ((iqh .eq. 0) .or. (iqh .eq. 2))then
c--24/10/91 generate matrix xmcv to allow individual noise
300      continue
          write(*,*)' Give Q limit to compute std: QMAX5(cfs)'
          read(*,*)qmax5
          write(*,6300)
6300     format(' Select:(0) Same CV to perturb all Qs.',
              /*,8x,'(1) Individual CV for each load.',
              /*,8x,'(2) Individual CV for each noad.',/)
          read(*,*)icv
          if (icv .eq. 0)then
              write(*,*)' Give CV to perturb all nodes : '
              read(*,*)cv
              do 310 i=1,nl
                  do 310 j=1,nn
                      xmcv(i,j)=cv
310          continue
              goto 390
          endif
          if (icv .eq. 1)then
              do 320 i=1,nl
                  write(*,*)' Give CV to perturb nodes from load ',i,':'
                  read(*,*)cv
                  do 320 j=1,nn
                      xmcv(i,j)=cv
320          continue
              goto 390
          endif
          if (icv .eq. 2)then
              do 330 i=1,nl
                  do 330 j=1,nn
                      write(*,*)' Give CV to perturb node ',j,' from load ',i,':'
                      read(*,*)cv
                      xmcv(i,j)=cv
330          continue
              goto 390
          endif
          write(*,*)' Invalid Selection '
          goto 300
390      continue
          write(*,*)' Condition noise on Q ? No(0), Yes(1) '
          read(*,*)icondq
          if(icondq .ne.1)icondq=0
          if(icondq .eq. 1)then
              write(*,*)' Give # of std to accept noise on Q '
              read(*,*)xnstdq
          endif
          endif
          if ((iqh .eq. 1) .or. (iqh .eq. 2))then
              write(*,*)' Option valid only when FGN is the last node'
              write(*,*)'      ***FGN will not be perturbed****'
7000     CONTINUE
          write(*,7300)
7300     format(' Select:(0) Same STD to perturb all Hs.',
              /*,8x,'(1) Individual STD for each load.',
              /*,8x,'(2) Individual STD for each noad.',/)
          read(*,*)istd
          if (istd .eq. 0)then
              write(*,*)' Give STD to perturb Hs from all nodes : '
              read(*,*)std
              do 7310 i=1,nl
                  do 7311 j=1,nn-1
                      xmstd(i,j)=std
7311          continue
                  xmstd(i,nn)=0.0
7310     continue
          endif

```

```

        goto 7390
    endif
    if (istd .eq. 1)then
        do 7320 i=1,nl
            write(*,*)' Give std to perturb Hs from load ',i,':'
            read(*,*)std
            do 7321 j=1,nn-1
                xmstd(i,j)=std
7321            continue
                xmstd(i,nn)=0.0
7320            continue
            goto 7390
        endif
        if (istd .eq. 2)then
            do 7330 i=1,nl
                do 7331 j=1,nn-1
                    write(*,*)' Give Std to perturb node',j,' from load ',i,':'
                    read(*,*)std
                    xmstd(i,j)=std
7331                continue
                    xmstd(i,nn)=0.0
7330                continue
            goto 7390
        endif
        write(*,*)' Invalid Selection '
        goto 7000
7390    continue

ccc        write(*,*) ' STD to perturb H (in feet)? '
ccc        read(*,*)std
c-----  Convert STD of pressure from psig to feet
          stdi=std/2.3076923
          write(*,*)' Condition noise on H ? No(0), Yes(1) '
          read(*,*)icondh
          if(icondh .ne.1)icondh=0
          if(icondh .eq. 1)then
              write(*,*)' Give # of std to accept noise on H '
              read(*,*)xnstdh
          endif
        endif
c----Define default variables for option iqh > 2
          icondr=0
          igap=5
c----Read simulation options for optios iqh=0,1 or 2
          if(iqh .lt. 3)then
              WRITE(*,*) ' NUMBER OF MONTE CARLO RUNS ?'
              READ (*,*)numpert
              IF (NUMPERT .GT. NPERT)then
                  write(*,*)' **WARNING: Maximum allowable is ',npert
                  NUMPERT=NPERT
              endif
              write(*,*) 'Give initial seed for random generator : '
              write(*,*) ' (Enter 0 for arbitrary seed ) '
              read(*,*)iseedjv
              write(*,*) 'Condition Monte Carlo for regressed Cs ?No(0),Yes(1):'
              read(*,*)icondr
              if(icondr .ne. 1)icondr=0
              if(icondr .eq.1)then
                  write(*,*)' Give lower bound to accept C '
                  read(*,*)clowb
                  write(*,*)' Give upper bound to accept C '
                  read(*,*)cupb
              endif
              WRITE(*,*) ' interval to compute statistics to check convergence '
              READ (*,*)igap
              if(igap .lt. 2)igap=2
              if(igap .gt. numpert)igap=numpert
c-----end of simulation options for iqh < 1
          endif

          WRITE(*,*) ' Save regressed Cs from each run ? NO(0), YES(1):'
          READ (*,*)ISAVEC
          if (isavec .gt. 0)then
              WRITE(*,*) ' Name of file to save Cs ?'
              READ (*,2)nfileoc

```

```

endif
c----Modified 17/8/90 to save H's perturbed
WRITE(*,*) ' Save Hs perturbed ? NO(0), YES(1):'
READ (*,*)ISAVEH
if (isaveh .gt. 0)then
    WRITE(*,*) ' Name of file to save Hs ?'
    READ (*,2)nfileoh
endif
7777 write(*,*)' Print Regression Matrix ? NO(0), YES(1):'
read(*,*)iprint
if ((numpert .gt. 2) .and. (iprint .gt.0))then
    write(*,*)' You must not print matrix X at each run'
    WRITE(*,*)' since you selected ',numpert,' runs'
    goto 7777
endif

if (isaveh .gt. 0)open(unit=nouh,file=nfileoh)
IF(IQH.GT.2)GOTO 31
19 continue
i=0
20 continue
i=i+1
c----*****This is limiting to drop at most 10 equations per load.*****
if(i .gt. 10)goto 30
WRITE(*,*)' DROP EQUAT. AT NODE # ?(0 to stop) '
c----Read from console
READ (*,*)NDROP(i)
if(ndrop(i) .eq. 0)goto 30
IF(NDROP(i) .GT. NN)THEN
    WRITE(*,*) ' TRYING TO DROP INVALID NODE ',NDROP,NN
    i=i-1
    GOTO 20
ENDIF
goto 20
30 continue
nundrop=i-1
31 CONTINUE
c----Test to see if is dropping at least one equation (redundant)
if (nundrop .eq. 0)then
    write(*,*)' At least one equation must be dropped '
    goto 19
endif
I=0
WRITE(NOUT,*)' LUMP PIPES : '
33 CONTINUE
I=I+1
IF(I .GT. NP)GOTO 35
WRITE(*,*)' Lump C at pipe ? (0 to stop): '
read(*,*)lvec(i)
if(lvec(i) .eq. 0)goto 35
if((lvec(i) .gt. np).or.(lvec(i) .lt. 0))then
    write(*,*)' Invalid pipe '
    i=i-1
    goto 33
endif
if(i.gt.1)then
    if(lvec(i) .le. lvec(i-1))then
        write(*,*)' Invalid pipe, try again '
        i=i-1
        goto 33
    endif
endif
WRITE(NOUT,*)' PIPE ',LVEC(I)
goto 33
35 continue
write(*,*)' Number of iterations to show time : '
read(*,*)nshow
c----Check dimension of arrays and solvability of system
IF ( NP .GT. ((NN-nundrop)*NL)) THEN
    WRITE(*,*)' Undetermined System ',NL,NN,nundrop,NP
    GO TO 9999
ENDIF
IF ( NL .GT. NUML) .OR. (NN .GT. NUMN) .OR. (NP .GT. NUMP))THEN
WRITE(*,*)' ARRAY OVERFLOW (Recompile) ',NL,NUML,NN,NUMN,NP,NUMP
GO TO 9999

```

```

ENDIF
write(*,*)' Print statistics matrices from scorr ? NO(0), YES(1)'
read(*,*)iprt
C-----Get physical characteristics of pipes and nodes from indata
ccc WRITE(*,*) ' GIVE GR. ELEV. FOR EACH NODE : '
ccc READ(NINP,*)(INODE(I,1),I=1,NN)
do 6 i=1,kn
6 inode(i,1)=e(i)
ccc WRITE(*,*) ' GIVE PIPE DIAM,LENGTH,NODE FROM, NODE TO'
DO 5 I=1,NP
ccc WRITE(*,*) ' PIPE ',I
ccc READ(NINP,*)PIPE(I,1),PIPE(I,2),IPIPE(I,1),IPIPE(I,2)
C-----Valid to convert to feet/inches
PIPE(I,1)=d(i)*12.
PIPE(I,2)=r(i)
IPIPE(I,1)=ja(i)
IPIPE(I,2)=jb(i)
C-----Get elevation of FGN
if(ja(i) .eq. 0)then
ipipe(i,1)=nn
inode(nn,1)=engy(i)
endif
5 CONTINUE
C-----Read node information for parameter estimation
WRITE(*,*) ' GIVE NODE FLOW AND PIEZ. PRESSURE (f10.5,f14.7): '
DO 10 K=1,NL
WRITE(*,*)' LOAD ',K,': '
DO 10 I=1,NN
READ(NINP,8801)XNODE(K,I,1),XNODE(K,I,2)
8801 format(f10.5,f14.7)
10 CONTINUE
write(*,*)' Computing constant for the pipes... '
C-----Identify pipes connected to a node
CALL SNCON(NP,NN,IPIPE,INODE)
C----Compute constant for pipes
CALL SPPC(NP,PIPE,PIPEK)
C*****
C PRINT VARIABLES
C*****
WRITE(NOUT,*)' WATER DISTRIBUTION NETWORK '
WRITE(NOUT,*)' *** Program Jvpipe4 ***'
WRITE(NOUT,*)' *Estimation of C using MLR*'
c*** OK for F77L MS-Fortran
ccc CALL DATE(DATEOUT)
ccc CALL TIME(TIMEOUT)
c---Ok for MS-Fortran
call getdat(iyr,imon,iday)
call gettim(ihr,imin,isec,i100th)
ccc WRITE(NOUT,*)' EXECUTED ON ',DATEOUT,' ',TIMEOUT
write(nout,8002)imon,iday,iyr,ihr,imin
8002 format(5x,'Executed on ',i2,'/',i2,'/',i4,2x,'at ',i2,':',i2)
WRITE(NOUT,1111)
1111 FORMAT(5X,'INPUT VARIABLES')
WRITE(NOUT,1009)
1009 FORMAT(4X,16('='))
WRITE(NOUT,1115)NL,NN,NP,(NDROP(i),i=1,ndrop)
1115 FORMAT(5X,'NUMBER OF LOADS, NL = ',I4/,
*5X,'NUMBER OF NODES, NN = ',I4/,
*5X,'NUMBER OF PIPES, NP = ',I4/,
*5X,'DROP EQUATION AT NODE(s) ',10I5)
WRITE(NOUT,1116)
1116 FORMAT(5X,'NODE DATA',/,5X,44('='))
WRITE(NOUT,*)' NODE ELEV. #PIPES CONNECTED PIPES'
DO 50 I=1,NN
WRITE(NOUT,5000)I,(INODE(I,J),J=1,8)
50 CONTINUE
WRITE(NOUT,1216)
1216 FORMAT(5X,44('='))
5000 FORMAT(5X,I3,4X,8(I3,2X))
WRITE(NOUT,*)' LOAD NODE FLOW TOTAL HEAD(ft)'
DO 55 K=1,NL
WRITE(NOUT,5006)k,(I,XNODE(K,I,1),XNODE(K,I,2),I=1,1)
55 WRITE(NOUT,5005)(I,XNODE(K,I,1),XNODE(K,I,2),I=2,NN)
5005 FORMAT(11X,I3,F12.3,2X,f14.7)
5006 format(5x,i3,3x,I3,F12.3,2X,f14.7)

```

```

WRITE(NOUT,1117)
1117 FORMAT(5X,'PIPE DATA',/,5X,53('='))
WRITE(NOUT,*)' PIPE DIAMETER LENGTH FROM TO K'
DO 60 I=1,NP
60 WRITE(NOUT,6001)I,PIPE(I,1),PIPE(I,2),IPIPE(I,1),IPIPE(I,2),
*PIPEK(I)
WRITE(NOUT,1217)
1217 FORMAT(5X,53('='))
6001 FORMAT(5X,I3,2X,2(F9.2,3X),2(I3,3X),E12.5)
WRITE(NOUT,*)'
c*****
c Decrease NN to account for at least 1 redundant equation
c Note: last node must be the FGN
c*****
cccccc NN=NN-1
c--The above statements is not needed since the user selects the nodes to drop
C----Compute coefficients for regression
      if (nundrop .eq. 0)then
          N=NL*NN
      else
          n=nl*(nn-nundrop)
      endif
      write(nout,1120)n
1120 format(5x,'NUMBER OF EQUATIONS, N= ',I5,/)
      write(nout,1121)nfilein,nfileou,iqh,deltah,deltaq,qmax5,icv
1121 format(5x,'Simulation options:',/5x,'nfilein ',a12,'nfileou ',a12,
*' iqh ',i1,' deltah ',f8.5,/,5x,' deltaq ',f8.5,' qmax5 ',f8.3,
8' icv ',i3)
      write(nout,1122)icondr,igap,isavec,isaveh,iprint,iprt,icondq,
*xnstdq,icondh,xnstdh
      if(icondr.eq.1)write(nout,1123)clowb,cupb
1122 format(5x,'icondr ',i1,' igap ',i3,' isavec ',i1,' isaveh ',i1,
*' iprint ',i1,' iprt ',i1,/,5x,'icondq ',i1,' xnstdq ',f7.3,
*' icondh ',i1,' xnstdh ',f7.3)
1123 format(25x,'clowb ',f10.5,' cupb ',f10.5)
      if(icv .eq. 0)write(nout,1024)icv,xmcv(1,1)
      if(icv .eq. 1)write(nout,1024)icv,(xmcv(i,1),i=1,nl)
      if(icv .eq. 2)write(nout,1024)icv,(xmcv(i,j),j=1,nn),i=1,nl)
1024 format(5x,'icv ',i2,' xmcv ',10(1x,f7.4))
      if(istd .eq. 0)write(nout,1025)istd,xmstd(1,1)
      if(istd .eq. 1)write(nout,1025)istd,(xmstd(i,1),i=1,nl)
      if(istd .eq. 2)write(nout,1025)istd,(xmstd(i,j),j=1,nn),i=1,nl)
1025 format(5x,'istd ',i2,' xmstd ',10(1x,f7.4))
      CALL SMAKEB(NL,NN,ndrop,XNODE,QJV,indlq,indnq)
      write(*,*)' Processing Simulations...'

C----Compute Sensitivity Matrix dC/dH----
      if((IQH .eq. 3) .or. (iqh .eq.5))then
          numpert=nl*(nn-1)+1
          i=0
          do 800 il=1,nl
              do 800 in=1,nn-1
                  i=i+1
                  write(*,*)' Regression # ',i, 'For dC/dH'
                  call sperth3(nl,nn,xnode,xnodep,isaveh,il,in,deltah)
                  CALL SMAKEX(NL,NN,NP,ndrop,XNODEP,INODE,IPIPE,PIPEK,X)
                  Run regression ON XNODEP for roughness coefficient "C"
                  CALL SMREGIA1(N,NP,X,QJV,CJV,0,iprint)
                  C---- Save values of C to compute statistics
                  call savec(np,i,csave,cjv)
800 CONTINUE
C----- Generate C's for base run(no noise)
                  CALL SMAKEX(NL,NN,NP,ndrop,XNODE,INODE,IPIPE,PIPEK,X)
                  Run regression ON XNODE for roughness coefficient "C"
                  CALL SMREGIA1(N,NP,X,QJV,CJV,0,iprint)
                  C---- Save values of C to compute statistics
                  i=i+1
                  call savec(np,i,csave,cjv)
                  if(numpert .ne. i)then
                      write(*,*)' Error numpert ne i ',numpert,i
                      goto 9999
                  endif
                  C---- Compute and save matrix dc/dh transpose
                  open(unit=ndcdh,file=fdcdh)
                  write(ndcdh,*)nl*kn,np

```

```

      do 801 j=1,nl*kn
      do 802 i=1,np
        dcdh(i)=(csave(j,i)-csave(numpert,i))/deltah
802      continue
        write(ndcdh,6040)(dcdh(i),i=1,np)
6040      format(10(e13.7,1x))
801      continue
      close(ndcdh)
      if(iqh .ne.5)goto 900
    endif

c----Compute Sensitivity Matrix dC/dQ
    if((IQH .eq. 4) .or. (iqh .eq.5))then
      numpert=nl*(nn-1)+1
      CALL SMAKEX(NL,NN,NP,ndrop,XNODE,INODE,IPIPE,PIPEK,X)
      DO 810 I=1,NUMPERT-1
        do 811 j=1,numpert-1
          qp(j)=qjv(j)
911          qp(i)=qjv(i)+deltaq
          write(*,*)' Regression # ',i, 'For dC/dQ'
          CALL SMREGIA1(N,NP,X,QP,CJV,0,iprint)
c----- Save values of C to compute statistics
          call savec(np,i,csave,cjv)
910        CONTINUE
c----- Generate C's for base run(no noise)
          CALL SMREGIA1(N,NP,X,QJV,CJV,0,iprint)
          call savec(np,i,csave,cjv)
c----- Save matrix dc/dq transpose
          open(unit=ndcdq,file=fdcdq)
          write(ndcdq,*)numpert-1,np
          do 812 j=1,numpert-1
            do 813 i=1,np
              dcdh(i)=(csave(j,i)-csave(numpert,i))/deltaq
913            continue
              write(ndcdq,6040)(dcdh(i),i=1,np)
912            continue
          close(ndcdq)
          goto 900
        endif

      IF (IQH .EQ. 1)THEN
c----- Introduce noise on H's
        DO 600 I=1,NUMPERT
c----- Re-set np to its original value at each iteration
          if((i-i/nshow*nshow) .eq. 0)then
            call gettim(ihr,imin,isec,i100th)
            WRITE(*,*)' REGRESSION # ',I,' time : ',ihr,imin
          endif
901        CONTINUE
          np=kk
          CALL SPERTH2(NL,NN,xmstd,XNODE,XNODEP,isaveh,icondh,xnsthd,
            *nviohl,nviohu)
          CALL SMAKEX(NL,NN,NP,ndrop,XNODEP,INODE,IPIPE,PIPEK,X)
          if (lvec(1) .gt. 0)then
            call slumpc(x,n,np,lvec,xl,nplump)
c----- Set np to the new value after lumping C's
            np=nplump
          else
            do 605 iig=1,n
              do 605 jjg=1,np
                xl(iig,jjg)=x(iig,jjg)
905            continue
          endif

c----- Run regression ON XNODEP for roughness coefficient "C"
          CALL SMREGIA1(N,NP,XL,QJV,CJV,0,iprint)
          IF(ICONDR .EQ. 1)THEN
            CALL SCONDR(NP,CJV,CLOWB,CUPB,NVIOCL,NVIOCU,IREJECT)
            IF(IREJECT .EQ. 1)GOTO 601
          ENDIF
c----- Save values of C to compute statistics
          call savec(np,i,csave,cjv)
900        CONTINUE
          GOTO 900
        ENDIF
      ENDIF

```



```

if (iqh .eq. 0)then
ccc   print *, ' Entering iqh = 0'
      CALL SMAKEX(NL,NN,NP,ndrop,XNODE,INODE,IPIPE,PIPEK,X)
ccc   print *, ' Smakex processed OK '
ccc   print *, ' Lvec(1) = ',lvec(1)
      if (lvec(1) .gt. 0)then
ccc     print *, ' Calling slumpc '
        call slumpc(x,n,np,lvec,xl,nplump)
ccc     print *, ' Slumpc processed OK. nplump = ',nplump
        np=nplump
ccc     print *, ' np dropped to ',np
      else
        do 105 i=1,n
          do 105 j=1,np
            xl(i,j)=x(i,j)
105      continue
        endif

C-----Loop to generate noise on Q
      DO 100 I=1,NUMPERT
        if((i-i/nshow*nshow) .eq. 0)then
          call gettim(ihr,imin,iseq,i100th)
          WRITE(*,*)' REGRESSION # ',I,' time : ',ihr,imin
        endif
101     CONTINUE
ccc     print *, ' Calling Spertq'
        CALL SPERTQ(N,XMCV,QJV,QP,indlq,indnq,qmax5,icondq,xnstdq,
          *nvioql,nvioqu)
C----- Run regression ON QP for roughness coefficient "C"
ccc     print *, ' Calling Smregia1'
        CALL SMREGIA1(N,NP,XL,QP,CJV,0,iprint)
c----- print *, ' After Calling smregia1
          IF(ICONDR .EQ. 1)THEN
            CALL SCONDR(NP,CJV,CLOWB,CUPB,NVIOCL,NVIOCU,IREJECT)
            IF(IREJECT .EQ. 1)GOTO 101
          ENDIF
C---- Save values of C to compute statistics
        call savec(np,i,csave,cjv)
100     CONTINUE
        goto 900
      endif

C----- Introduce noise on H's and Q's
      if (iqh .eq. 2)then
        DO 700 I=1,NUMPERT
          if((i-i/nshow*nshow) .eq. 0)then
            call gettim(ihr,imin,iseq,i100th)
            WRITE(*,*)' REGRESSION # ',I,' time : ',ihr,imin
          endif
701     CONTINUE
ccc     CALL SPERTH2(NL,NN,std,XNODE,XNODEP, isaveh)
        CALL SPERTH2(NL,NN,xmstd,XNODE,XNODEP, isaveh, icondh, xnstdh,
          *nviohl,nviohu)
        np=kk
        CALL SMAKEX(NL,NN,NP,ndrop,XNODEP,INODE,IPIPE,PIPEK,X)
        if (lvec(1) .gt. 0)then
          call slumpc(x,n,np,lvec,xl,nplump)
          np=nplump
        else
          do 705 iig=1,n
            do 705 jjg=1,np
              xl(iig,jjg)=x(iig,jjg)
705      continue
          endif
cccccc CALL SPERTQ(N,XMCV,QJV,QP,indlq,indnq)
        CALL SPERTQ(N,XMCV,QJV,QP,indlq,indnq,qmax5,icondq,xnstdq,
          *nvioql,nvioqu)

C----- Run regression ON XNODEP and Qp for roughness coefficient "C"
        CALL SMREGIA1(N,NP,XL,QP,CJV,0,iprint)
        IF(ICONDR .EQ. 1)THEN
          CALL SCONDR(NP,CJV,CLOWB,CUPB,NVIOCL,NVIOCU,IREJECT)
          IF(IREJECT .EQ. 1)GOTO 701
        ENDIF
C----- Save values of C to compute statistics

```

```

        call savec(np,i,csave,cjv)
700    CONTINUE
    endif
c ----- Write C's for posterior statistics
900    CONTINUE
    IF (isavec .gt. 0)then
        open (unit=nouc,file=nfileoc)
        write(nouc,*)np,',',numpert
        write(nouc,2021)(I,I=1,NP)
c-----format ready for 10 pipes
2021    FORMAT(20('PIP',I2,','))
        do 200 i=1,numpert
cccccccc        write(nouc,6000)i,(csave(i,j),j=1,np)
cccc 6000        format(5x,i3,10(f7.2,1x))
                write(nouc,6010)(csave(i,j),j=1,np)
6010            format(10(f9.5,1x))
200            continue
                close(nouc)
        ENDIF

        WRITE(*,*)TITLE
        WRITE(NOUT,1110)TITLE
1110    FORMAT(5X,A70,/)
c-----Compute statistics on C's

        CALL SSTAT(NUMPERT,NP,CSAVE,CSTAT)
        write(nout,*)' # regressions noise on CV for Q STD for H'
        write(nout,1090)numpert,choice(iqh+1),CV,stdi,std
1090    format(2x,i5,13x,a5,2x,f7.3,2x,f9.5,2x,' psig',f7.3,' ft',/)
        write(nout,*)' STATISTICS OF C VALUES '
                write(nout,6002)(i,i=1,np)
6002    format(1x'Pipe # ',11(2x,i6,2x))
6004    format(1x,a6,1x,11(f9.3,1x))
        i=1
        write(nout,6004)nsta(i),(cstat(i,j),j=1,np)
        i=2
        write(nout,6004)nsta(i),(sqrt(cstat(i,j)),j=1,np)
        if(iprint .eq. 0)call scorr(csave,numpert,np,nsave,fsave,iprt)
c---Save statistics to check convergence
        OPEN(UNIT=NSAVEM,FILE=FSAVEM)
        OPEN(UNIT=NSAVEST,FILE=FSAVEST)
        DO 8000 I=IGAP,NUMPERT,IGAP
            CALL SSTAT(I,NP,CSAVE,CSTAT)
            WRITE(NSAVEM,6700)I,(CStat(1,J),J=1,NP)
            WRITE(NSAVEST,6800)I,(dsqrt(CStat(2,J)),J=1,NP)
8000    CONTINUE
        IF((I-IGAP) .NE. NUMPERT)THEN
            CALL SSTAT(NUMPERT,NP,CSAVE,CSTAT)
            WRITE(NSAVEM,6700)NUMPERT,(CStat(1,J),J=1,NP)
            WRITE(NSAVEST,6800)NUMPERT,(dsqrt(CStat(2,J)),J=1,NP)
        ENDIF
        CLOSE(NSAVEM)
        CLOSE(NSAVEST)
6700    FORMAT(i5,16(1x,f13.6))
6800    FORMAT(i5,16(1x,f13.8))
c-----Write statistics of violations if condition MC
        write(nout,1124)nviocl,nviocu,nvioql,nvioqu,nviohl,nviohu
1124    format(5x,'Statistics on condition MC',/,5x,'nviocl ',i7,
        *' nviocu ',i7,' nvioql ',i7,' nvioqu ',i7,' nviohl ',i7,
        *' nviohu ',i7)
9999    continue
        WRITE(*,*)' (NORMAL END OF PROGRAM "JVPIPE4")'
        CLOSE(NINP)
        CLOSE(NOUT)
        STOP
        END

SUBROUTINE SPPC(NP,PIPE,PIPEK)
C *****
C COMPUTES CONSTANT K1 - PHYSICAL PARAMETER FOR PIPES
C *****
IMPLICIT REAL*8(A-H,O-Z)
INTEGER NUMP,NUMN
PARAMETER (NUMP=51,NUMN=51)

```

```

      DIMENSION PIPE(NUMP,2),PIPEK(NUMP)
C----Define constants for: Q=cfs, H=ft
C-----C1=2.63, C2=0.54-----
      C1=4.87/1.852
      C2=1./1.852
c -----modified 20-08-90      C3=2.3136
      c3=2.3123067
C----Pressure in feet of water, diameter in inches
      DO 10 I=1,NP
          PIPEK(I)=((PIPE(I,1)/12.)**C1)/(PIPE(I,2)**C2)/C3
10  CONTINUE
      RETURN
      END

      SUBROUTINE SMAKEB(NL,NN,ndrop,XNODE,B,indlb,indnb)
c *****
c Generates constant vector B for regression
c (external demands)
c *****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER NUMP,NUMN
      PARAMETER (LIMIT=70,NUMP=51,NUMN=51,NUML=5)
      DIMENSION XNODE(NUML,NUMN,2),B(LIMIT),ndrop(10),
*indlb(limit),indnb(limit)
c-- Generate vectors b, indlb and indnb
      do 5 i=1,limit
          b(i)=0.0
          indlb(i)=0
          indnb(i)=0
5  continue
C----Generate constant vector B (External flow at nodes)
      INC=1
      DO 10 K=1,NL
          icont=1
          DO 10 J=1,NN
c----- Drop equation at node NDRROP
              if(j .eq. ndrop(icont))then
                  icont=icont+1
                  goto 9
              endif
              B(INC)=XNODE(K,J,1)
c-- 24/10/91 save load and node corresponding to each element of B
              indlb(inc)=k
              indnb(inc)=j
              INC=INC+1
9  continue
10  CONTINUE
      RETURN
      END

      SUBROUTINE SNCON(NP,NN,IPIPE,INODE)
c *****
c * SUBROUTINE: SNCON (Node Connections) *
c * * * * *
c * REMARKS : Identifies pipes connected to a node *
c * * * * *
c * VARIABLES *
c * ===== *
c * NP : # OF PIPES *
c * NN : # OF NODES *
c * IPIPE : MATRIX (from node, to node) *
c * INODE : MATRIX (G.E.,# conn,pip1,pip2,...) *
c * IND(i) : STORE next available position *
c * NODEN : Get node number *
c *****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER NUMP,NUMN,NUML
      PARAMETER (NUMP=51,NUMN=51,NUML=5)
      DIMENSION INODE(NUMN,8),IPIPE(NUMP,2),IND(NUMN)
C----Initialize index for next free position
      DO 10 I=1,NN
10  IND(I)=0
C----Generate node connections
      DO 20 I=1,NP
          DO 20 J=1,2

```

```

      NODEN=IPIPE(I,J)
      INODE(NODEN,IND(NODEN)+3)=I
      IND(NODEN)=IND(NODEN)+1
      INODE(NODEN,2)=IND(NODEN)
20  CONTINUE
      RETURN
      END

SUBROUTINE SMAKEX(NL,NN,NP,ndrop,XNODE,INODE,IPIPE,PIPEK,X)
*****
C      * SUBROUTINE: SMAKEX *
C      * *
C      * REMARKS : Generates coefficient matrix for regres *
C      * sion. *
C      * *
C      * VARIABLES *
C      * ===== *
C      * NP : # OF PIPES *
C      * NN : # OF NODES *
C      * NL : # OF LOADS *
C      * ndrop(i) : drop equation at this node *
C      * IPIPE : MATRIX (from node, to node) *
C      * INODE : MATRIX (G.E.,# conn,pip1,pip2,...) *
C      * XNODE : MATRIX (Load, Flow, Head) *
C      * X : Matrix (Coeffic. for regression) *
C      * INC : Row index *
C      * HI : Pressure head at node i (from) *
C      * HJ : Pressure head at node j (to) *
C      * PIPEN : Get pipe number *
C      * NC : # of pipes connected to the node *
C      * SMULT : -1 flow is leaving node *
C      * 1 flow is entering node *
*****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER NUMP,NUMN,NUML,PIPEN
      PARAMETER (LIMIT=70,NUMP=51,NUMN=51,NUML=5)
      DIMENSION XNODE(NUML,NUMN,2),INODE(NUMN,8),IPIPE(NUMP,2),
      *X(LIMIT,LIMIT),PIPEK(NUMP),ndrop(10)
      C2=1./1.852
C-----Initialize matrix X
      DO 10 I=1,NL*NN
      DO 10 J=1,NP
          X(I,J)=0.0
10  CONTINUE
      INC=0
      DO 100 K=1,NL
      icont=1
      DO 100 I=1,NN
C----- Drop equation at node NDROP
          if (i .eq. ndrop(icont))then
              icont=icont+1
              goto 90
          endif
          INC=INC+1
          NC=INODE(I,2)
C-----Loop for each pipe connecting to node I
          DO 200 J=1,NC
              PIPEN=INODE(I,J+2)
C--- Node in question is I
              NODEN=IPIPE(PIPEN,1)
C----- Get the other extreme of the pipe
              if (noden .eq. i)NODEN=IPIPE(PIPEN,2)
              HI=XNODE(K,NODEN,2)
              HJ=XNODE(K,I,2)
              AUX=DABS(HI-HJ)**C2*PIPEK(PIPEN)
              X(INC,PIPEN)=DSIGN(AUX,HI-HJ)
200  CONTINUE
          90  continue
100  CONTINUE
C----- display matrix X
      write(6,*)' Matrix X'
      call sprtmat(x,inc,np,limit)
      ni=0
      do 900 i=1,inc
          ni=ni+1
      c

```

```

c          if (ni .gt. nn)then
c              ni=1
c              ln=(i+nn-1)/nn
c              write(6,*)' Load ',ln
c          endif
c          write(6,1000)ni,(x(i,j),j=1,np)
c 900      continue
c 1000    format(1x,I1,10(1x,e14.6))
          RETURN
          END

          SUBROUTINE SSTAT(NUM, NP, XMAT, VECT)
          *****
          C          * SUBROUTINE: SSTAT
          C          *
          C          * REMARKS : COMPUTES MEAN AND VARIANCE OF VALUES
          C          *
          C          * FROM QP AND C
          C          *
          C          * VARIABLES
          C          *
          C          * =====
          C          *
          C          * NP : # OF PIPES
          C          *
          C          * NUM : # PERTURBATIONS
          C          *
          C          * XMAT : MATRIX (NUM, NP)
          C          *
          C          * (CONTAINS VALUES OF COEF. FOR EACH
          C          *
          C          * PERTURBATIONS)
          C          *
          C          * VECT : SAVES STATISTICS FOR XMAT
          C          *
          C          * VECT(1, J)=MEAN OF XMAT(I, J)
          C          *
          C          * VECT(2, J)=VARIANCE OF XMAT(I, J)
          C          *
          C          * SUM(J): SAVE PARTIAL RESULTS
          C          *
          C          *
          C          *****
          IMPLICIT REAL*8(A-H, O-Z)
          INTEGER NP, NUM
          PARAMETER (LIMIT=70, NUMP=51, NUMN=51, NUML=5, NPRT=1000)
          DIMENSION XMAT(NPRT, 20), VECT(2, 20), SUM(20)
          IF (NUM .GT. NPRT) THEN
              WRITE(*, *) '# OF PERT. EXCEEDS DIMENSION OF XMAT ', NUM, NPRT
              GOTO 999
          ENDIF
          DO 10 J=1, NP
              SUM(J)=0.0
          10  continue
          c-----computes mean
          DO 100 I=1, NUM
              DO 100 J=1, NP
                  SUM(J)=SUM(J)+XMAT(I, J)
          100 CONTINUE
          DO 110 J=1, NP
              VECT(1, J)=SUM(J)/(NUM)
              SUM(J)=0.0
          110 CONTINUE
          c-----computes variance
          DO 200 I=1, NUM
              DO 200 J=1, NP
                  SUM(J)=SUM(J)+(XMAT(I, J)-VECT(1, J))*(XMAT(I, J)-VECT(1, J))
          200 CONTINUE
          DO 210 J=1, NP
              VECT(2, J)=SUM(J)/(NUM-1)
          210 CONTINUE
          ccc  WRITE(*, *)' <NORMAL END SSTAT>'
          999  RETURN
          END

          SUBROUTINE SPERTQ(N, XMCV, Q, QP, indlq, indnq, qmax5, icondq, xnstdq,
          *nvioql, nvioqu)
          *****
          C          * SUBROUTINE: SPERQ
          C          *
          C          * REMARKS : GENERATES NORMALLY DISTRIBUTED NOISE
          C          *
          C          * AND PERTURB Q VECTOR
          C          *
          C          *
          C          * VARIABLES
          C          *
          C          * =====
          C          *
          C          * N : # OF NODES
          C          *
          C          *****

```

```

C      *          Q : VECTOR OF ORIGINAL VALUES      *
C      *          QP : CONTAINS Q+NOISE                *
C      *          xmcv(l,n) : contains individual cv for each node *
C      *          CV : DESIRED COEFFICIENT OF VARIATION *
C      *          indlq(i) : load of the ith element of Q      *
C      *          indnq(i) : noad of the ith element of Q      *
C      *          STD : STANDARD DEVIATION OF NOISE          *
C      *          = CV * Q(I)                             *
C      *          QMAX5 : LIMIT TO GENERATE MAXIMUM STD OF ERR.*
C      *          icondq : 1-Condition noise on Q;0-Not cond. *
C      *          xnstdq : # of std to reject noise          *
C      *          nvioql,nvioqu : # of violations upper and lower *
C      *          *****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER N
      PARAMETER (LIMIT=70,NUMP=51,NUMN=51,NUML=5)
      DIMENSION Q(LIMIT),QP(LIMIT),xmcv(numl,numn),indlq(limit),
      *indnq(limit)
C-----GENERATE NOISE MEAN 0, STD=CV*Q
      DO 10 I=1,N
C--      24/10/91 retrieve individual CV for each node and load
          iload=indlq(i)
          inoad=indnq(i)
          cv=xmcv(iload,inoad)
          STD=CV*Q(I)
          IF (Q(I) .GT. QMAX5)STD=CV*QMAX5
      5      XNOISE=GASDEV(I)
C--- Test of condition MC
          IF(ICONDQ .EQ. 1)THEN
              if(xnoise .lt. (-1.0*xnstdq))then
                  nvioql=nvioql+1
                  goto 5
              endif
              if(xnoise .gt. xnstdq)then
                  nvioqu=nvioqu+1
                  goto 5
              endif
          ENDIF
          QP(I)=Q(I)+XNOISE*STD
      10 CONTINUE
Ccc write(*,*)' Normal end spertq'
      RETURN
      END

      SUBROUTINE SPERTH2(NL,NN,XMSTD,XNODE,XNODEP,isaveh,icondh,xnstdh,
      *nviohl,nviohu)
C      *****
C      * SUBROUTINE: SPERTH2 (Aug 2,1990) *
C      * REMARKS : GENERATES NORMALLY DISTRIBUTED NOISE *
C      * AND PERTURB H VALUES by a desired STD *
C      * 6/8/90 modified *
C      * VARIABLES *
C      * ===== *
C      * NL : # OF LOADS *
C      * NN : # OF NODES *
C      * XNODE(NL,NN,2): MATRIX CONTAINING Q's & H's *
C      * XNODEP(NL,NN,2):NEW MATRIX WITH NOISE IN H's *
C      * XMSTD(I,J) : STD OF Hj from load I in ft *
C      * icondh : 1=condition noise on H; 0=not cond. *
C      * xnstdh : # of std to condition reject noise *
C      * nviohl,nviohu : statistics on the violations of noise*
C      * *****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER NL,NN
      PARAMETER (LIMIT=70,NUMP=51,NUMN=51,NUML=5)
      DIMENSION XNODE(NUML,NUMN,2),XNODEP(NUML,NUMN,2),xmstd(numl,numn)
      COMMON /jvga/NOU,nouh
C-----GENERATE NOISE MEAN 0, Standard Deviation =STD AND PERTURB H's
      DO 10 K=1,NL
          DO 20 I=1,nn-1
      21      XNOISE=GASDEV(I)
              IF(ICONDH .EQ. 1)THEN
                  if(xnoise .lt. (-1.0*xnstdh))then
                      nviohl=nviohl+1
                      goto 21
                  endif
              endif
          END DO
      END DO

```

```

        endif
        if(xnoise .gt. xnstdh)then
            nviohu=nviohu+1
            goto 21
        endif
    ENDIF
    std=xmstd(k,i)
    XNODEP(K,I,2)=XNODE(K,I,2)+XNOISE*STD
C----- Check if Hpert. is greater than the FGN:H(nn)
C----- Modified 6/8/90, 11/26/91
        if (XNODEP(K,I,2) .GT. XNODE(K,NN,2))then
            write(*,*)' invalid noise on node ',i,' load ',k
            goto 21
        endif
        XNODEP(K,I,1)=XNODE(K,I,1)
    20 continue
C----- Node nn is a FGN and is not being perturbed
        XNODEP(K,nn,2)=XNODE(K,nn,2)
        XNODEP(K,nn,1)=XNODE(K,nn,1)
    10 CONTINUE
C----- modified 17/08/90 to save H's for statistics
        if (isaveh .gt. 0)then
            do 100 k=1,nl
                write(nouh,8000)k,(xnodep(k,i,2),i=1,nn)
    100 continue
            endif
    8000 format(1x,i2,1x,20(f9.5,1x))
        RETURN
    END

SUBROUTINE SPERTH3(NL,NN,XNODE,XNODEP,isaveh,IL,IN,DELTAH)
C *****
C * SUBROUTINE: SPERTH3 (May 9, 1991) *
C * REMARKS : PERTURB H'S BY DELTAH TO COMPUTE dc/dH *
C * VARIABLES *
C * ===== *
C * NL : # OF LOADS *
C * NN : # OF NODES *
C * XNODE(NL,NN,2): MATRIX CONTAINING Q's & H's *
C * XNODEP(NL,NN,2):NEW MATRIX WITH NOISE IN H's *
C * DELTAH : PERTURBATION ON H'S (IN FEET) *
C *****
IMPLICIT REAL*8(A-H,O-Z)
INTEGER NL,NN
PARAMETER (LIMIT=70,NUMP=51,NUMN=51,NUML=5)
DIMENSION XNODE(NUML,NUMN,2),XNODEP(NUML,NUMN,2)
COMMON /jvga/NOU,nouh
DO 10 K=1,NL
DO 10 I=1,nn
    XNODEP(K,I,2)=XNODE(K,I,2)
    XNODEP(K,I,1)=XNODE(K,I,1)
    10 continue
XNODEP(IL,IN,2)=XNODE(IL,IN,2)+DELTAH
if (isaveh .gt. 0)then
    do 100 k=1,nl
        write(nouh,8000)k,(xnodep(k,i,2),i=1,nn)
    100 continue
endif
    8000 format(1x,i2,1x,20(f9.5,1x))
    RETURN
    END

SUBROUTINE SMREGIA1(N,NP,X,Y,B,ICONS,iprint)
C *****
C PERFORM MULTIPLE LINEAR REGRESSION
C *****
IMPLICIT REAL*8(A-H,O-Z)
REAL*8 MSE,MSR
INTEGER LIMIT
PARAMETER (LIMIT=70)
DIMENSION X(LIMIT,LIMIT),Y(LIMIT),XT(LIMIT,LIMIT),
*XTX(LIMIT,LIMIT),XTY(LIMIT,1),B(LIMIT),XTXINV(LIMIT,LIMIT),
C----- *bt(limit,limit),yt(limit,limit),yty(limit),btxy(limit),
*covmat(limit,limit),cormat(limit,limit),ypred(limit),e(limit),
*xmean(limit)

```

```

COMMON /jvga/NOU,nouh
c--Write initial values to check regression
CCC   print *, ' SMREGIA1: Entering Smregia1'
      if (iprint .gt. 0)then
        write(nout,*)' Subroutine Smregia-Run regression'
        write(nout,*)' N = ',n,' NP = ',np
        write(nout,*)' Matrix X used for regression'
        do 5 i=1,n
          write(nout,6)(x(i,j),j=1,np)
6       format(11(e14.6,1x))
5       continue
        write(nout,*)' Constant vector Y'
        write(nout,6)(y(j),j=1,n)
      endif
C-----Check range for ICONS (repeated below)
ccc   IF ( ICONS .LT. 0) ICONS=0
ccc   IF ( ICONS .GT. 1) ICONS=1
ccc   NP1=NP+ICONS
c-----Compute mean of observations
      YMEAN=0.0
      DO 10 J=1,NP
10     XMEAN(J)=0.0
      DO 20 I=1,N
        YMEAN=YMEAN+Y(I)
        DO 20 J=1,NP
20     XMEAN(J)=XMEAN(J)+X(I,J)
      YMEAN=YMEAN/N
      DO 30 J=1,NP
30     XMEAN(J)=XMEAN(J)/N

C-----Compute covariance matrix
      DO 40 I=1,N
40     X(I,NP+1)=Y(I)
      DO 45 I=1,NP+1
        DO 45 J=1,NP+1
45     COVMAT(I,J)=0.0
      XMEAN(NP+1)=YMEAN
      DO 46 K=1,NP+1
        DO 46 J=1,K
          SUM=0.0
          DO 47 I=1,N
47     SUM=SUM+(X(I,K)-XMEAN(K))*(X(I,J)-XMEAN(J))
          COVMAT(K,J)=SUM/float(n-1)
46     CONTINUE
CCC   write(nout,*) ' '
CCC   write(nout,*) '   Covariance matrix'
CCC   write(nout,*) '   ====='
CCC   call sprtmat(covmat,np+1,np+1,limit)
CCC   PRINT *, 'SMREGIA1: Compute correlation matrix'
      DO 49 I=1,NP+1
        DO 49 J=1,I
CCC   print *, 'Generating Cormat: i/j ',i,j
CCC   print *, ' Covmat(i,i)=' ,covmat(i,i),covmat(j,j)
49   CORMAT(I,J)=COVMAT(I,J)/SQRT(COVMAT(I,I))/SQRT(COVMAT(J,J))
      if(iprint .gt. 0)then
        write(nout,*) ' '
        write(nout,*) '   Correlation matrix'
        write(nout,*) '   ====='
        call sprtmat(cormat,np+1,np+1,limit)
        write(nout,*) ' '
      endif
C----Make augmented matrix X = (1,xi1,xi2,...xinp)
      IF ( ICONS .LT. 0) ICONS=0
      IF ( ICONS .GT. 1) ICONS=1
      NP1=NP+ICONS
      IF ( ICONS .Eq. 1) CALL SAUGMAT(X,N,NP)
C----Compute transpose of matrix X
ccc   print *, 'SMREGIA1: Calling Stranmat'
      CALL STRANMAT(X,N,NP1,XT)

C----Compute products

      CALL SMULTMAT(XT,NP1,N,X,N,NP1,XTX)
      if(iprint .gt. 0)then
        write(nout,*)'   MATRIX (XtX) '

```



```

        write(nout,*)' ===== '
        call sprtmat(xtx,np1,np1,limit)
    endif
    CALL SINVMAT(NP1,XTX,XTXINV)
    if(iprint .gt. 0)then
        WRITE(NOUT,*)'      MATRIX (XtX) Inverse '
        write(nout,*)' ===== '
        call sprtmat(xtxinv,np1,np1,limit)
    endif
    CALL SMULTMAT(XT,NP1,N,Y,N,1,XTY)
C           write(nout,*)'      VECTOR XtY'
C           write(nout,*)' ====='
C           call sprtmat(xty,np1,1,limit)
C-----Compute coefficients of regression b
    call smultmat(xtxinv,np1,np1,xy,np1,1,b)
C           write(nout,*)'      VECTOR b'
C           write(nout,*)' ====='
C           call sprtmat(b,np1,1,limit)
C----Calculate Ypred, E(Residuals),SSR,SSE,SSTO
    SSR=0.0
    SSE=0.0
    SSTO=0.0
    DO 50 I=1,N
        YPRED(I)=0.0
        DO 60 J=1,NP1
60           YPRED(I)=YPRED(I)+B(J)*X(I,J)
            E(I)=Y(I)-YPRED(I)
            SSR=SSR+(YPRED(I)-YMEAN)*(YPRED(I)-YMEAN)
            SSE=SSE+E(I)*E(I)
            SSTO=SSTO+(Y(I)-YMEAN)*(Y(I)-YMEAN)
50        CONTINUE

C-----Compute MSR, MSE, S, F, RR
C           MSR=SSR/(NP)
C           MSE=SSE/(N-NP1)
C           S=SQRT(MSE)
C           F=MSR/MSE
C           RR=SSR/SSTO

C           write(nout,1000)RR,S
C 1000 FORMAT(1H,5X,'MULTIPLE COEFF. OF DETERMINATION, R^2 = ',f12.3,
C           */,5X,' STANDARD ERROR OF ESTIMATE, S = ',f12.3)
C           WRITE(NOUT,1005)
C 1005 FORMAT(/,5X,'VARIABLE',3X,'COEFFICIENT',3X,'VARIANCE ',3X,
C           *'STD. ERROR',6X,'T')
C           WRITE(NOUT,1006)
C 1006 FORMAT(5X,65(' '))
C           DO 200 I=1,NP1
C               VAR=MSE*XTXINV(I,I)
C               SB=SQRT(VAR)
C               T=B(I)/SB
C               WRITE(NOUT,1007)I-ICONS,B(I),VAR,SB,T
C 1007       FORMAT(7X,'b(',I2,')',4X,f12.3,1X,f12.3,1X,f12.3,3X,f12.3)
200           CONTINUE
C           WRITE(NOUT,1006)

C           WRITE(NOUT,1010)
C 1010 FORMAT(/,15X,'ANALYSIS OF VARIANCE - ANOVA ',/,
C           *5X,'SOURCE          D.F.          SUM-OF-SQUARES          MEAN-SQUARE')
C           WRITE(NOUT,1011)NP,SSR,MSR,N-NP1,SSE,MSE,N-ICONS,SSTO
C 1011 FORMAT(5X,'REGRESSION          ',I2,6X,f12.3,7X,f12.3,/,
C           *5X,'ERROR          ',I2,6X,f12.3,7X,f12.3,/,
C           *5X,'TOTAL          ',I2,6X,f12.3,/)
C           WRITE(NOUT,*)'          SSTO-(SSR+SSE) ',SSTO-SSR-SSE
C           WRITE(NOUT,1015)F
C 1015 FORMAT(/,'          F-RATIO = ',f12.3)
C           write(nout,2000)
C 2000 format(/,5X,' ESTIMATE          RESIDUAL')
C           DO 300 I=1,N
C 300           WRITE(NOUT,2020)YPRED(I),E(I)
C 2020 FORMAT(6X,F9.3,6X,F9.3)
9999 continue

CCC       WRITE(*,*)' (NORMAL END OF SUBROUTINE "Smregia1")'
        RETURN

```

```

END

SUBROUTINE Scorr(XX,N,NP,nsave,fsave,iprt)

c *****
c Compute statistics on regressed values and save special
c file with np,meanC, stdC, covmat C, trace (covmat C), corrmatrix,
c and np.
c *****
IMPLICIT REAL*8(A-H,O-Z)
INTEGER LIMIT
PARAMETER (npert=1000,limit=70,nump=51)
DIMENSION XX(npert,20),covmat(nump,nump),cormat(nump,nump),
*xmean(nump),factor(nump,nump)
COMMON /jvga/NOUT,nouh
character*20 fsave
c---- variable iprt=0 suppress output of results from scorr in unit nout
c-----Compute mean of observations
DO 10 J=1,NP
  10 XMEAN(J)=0.0
  DO 20 I=1,N
    DO 20 J=1,NP
      20 XMEAN(J)=XMEAN(J)+XX(I,J)
  DO 30 J=1,NP
    30 XMEAN(J)=XMEAN(J)/float(N)
    if(iprt .ne. 0)then
      write(nout,*) ' '
      write(nout,*) ' Values computed and printed by scorr '
      write(nout,6004)(xmean(j),j=1,np)
6004 format(1x,'mean ',10(f13.6,1x))
      write(nout,*) ' '
    endif
c-----Compute covariance matrix
DO 45 I=1,Np
DO 45 J=1,Np
  45 COVMAT(I,J)=0.0
DO 46 K=1,NP
DO 46 J=1,K
  SUM=0.0
DO 47 I=1,N
  47 SUM=SUM+(XX(I,K)-XMEAN(K))*(XX(I,J)-XMEAN(J))
COVMAT(K,J)=SUM/float(n-1)
46 CONTINUE

  if(iprt .ne. 0)then
    write(nout,6005)(sqrt(covmat(j,j)),j=1,np)
6005 format(1x,'std ',10(f13.8,1x))
    write(nout,*) ' '
    write(nout,*) ' Covariance matrix'
    write(nout,*) ' ====='
    call sprtmat(covmat,np,np,nump)
  endif
c----Compute Trace of covmat
trace=0.0
do 100 i=1,np
  trace=trace+covmat(i,i)
100 continue
  write(nout,6008)trace
6008 format(/,1x,' Trace of COV C = ',e13.7)
c-----Compute correlation matrix
DO 49 I=1,NP
DO 49 J=1,I
  49 CORMAT(I,J)=COVMAT(I,J)/SQRT(COVMAT(I,I))/SQRT(COVMAT(J,J))
  if(iprt .ne. 0)then
    write(nout,*) ' '
    write(nout,*) ' Correlation matrix'
    write(nout,*) ' ====='
    call sprtmat(cormat,np,np,nump)
    write(nout,*) ' '
  endif
  write(*,*)' Calling sfload'
c----Compute component loadings
call sfload(cormat,NP,NUMP,factor,iprt)
write(*,*)' Saving results for later use '
c----Save statistics for later use: file=for066.dat

```

```

        open(unit=nsave,file=fsave)
        write(nsave,6010)np
6010 format(i5)
        write(nsave,6020)(xmean(j),j=1,np)
6020 format(10(f13.6,1x))
        write(nsave,6030)(dsqrt(covmat(j,j)),j=1,np)
6030 format(10(f13.8,1x))
        do 105 i=1,np
            write(nsave,6040)(covmat(i,j),j=1,i)
105 continue
        write(nsave,6040)trace
6040 format(10(e13.7,1x))
        do 110 i=1,np
            write(nsave,6040)(cormat(i,j),j=1,i)
110 continue
        do 120 i=1,np
            write(nsave,6040)(factor(i,j),j=1,np)
120 continue
        write(nsave,6010)np
        close(nsave)
ccc write(*,*)' Normal end Scorr'
        RETURN
        END

        SUBROUTINE SAVEC(NP,I,CSAVE,C)
C *****
C SAVE ESTIMATES OF C FOR ANALYSIS
C *****
        IMPLICIT REAL*8(A-H,O-Z)
        PARAMETER (LIMIT=70,NPERT=1000)
        DIMENSION C(LIMIT),CSAVE(NPERT,20)
        COMMON /jvga/NOUT,nouh
        DO 100 J=1,NP
            CSAVE(I,J)=C(J)
            if(c(j) .lt. 0.0)write(nout,*)' C=',c(j),' at pipe ',j,
* in run ',i
100 CONTINUE
        return
        end

        SUBROUTINE SFLOAD(CORR,NPIPES,NMAX,A,iprt)
C *****
C * SUBROUTINE: SFLOAD *
C * * * * *
C * REMARKS : Computes the component loadings of the *
C * Correlation matrix of regressed *
C * parameters C *
C * * * * *
C * VARIABLES *
C * ===== *
C * CORR : CORRELATION MATRIX OF C SHOWING ONLY *
C * LOWER TRIANGULAR ELEMENTS *
C * NPIPES : DIMENSION OF MATRIX CORR & A *
C * NMAX : PHYSICAL DIMENSION OF MATRIX CORR & A *
C * A : COMPONENT LOADINGS OF CORR: *
C * Sqrt(ALPHA(J)) * A(I,J) Where: *
C * alpha=eigenvalues; a=eigenvectors *
C * IPRT : = 0 : SUPPRESS OUTPUT TO UNIT NOUT *
C *****

        implicit real*8(a-h,o-z)
        PARAMETER(LIMIT=70,NUMP=51)
        DIMENSION A(NMAX,NMAX),D(Nump),E(Nump),CORR(NMAX,NMAX)
        common /jvga/nout
C CHECK COMPATIBILITY OF ARRAY DIMENSION
        IF(NMAX .NE. NUMP)THEN
            WRITE(NOUT,*)' ***** /
            WRITE(NOUT,*)' CHECK ARRAY DIMENSIONS IN SUBROUTINE SFLOAD:'
            WRITE(NOUT,*)' NMAX = ',NMAX,' NUMP = ',NUMP
            WRITE(NOUT,*)' ***** /
        ENDIF
c-----Create working matrix A
        do 10 i=1,npipes
            do 10 j=1,i
                a(i,j)=corr(i,j)

```

```

10  continue
c----Rename variables for eigen subroutines
n=npipes
np=ncmp
C---- set limit to round eigenvalue to zero
zerolim=0.0000001
write(nout,*) '   zerolim to set eigenvalues to zero is: ',
*zerolim

c--- fill the upper diagonal of A
do 20 i=1,n-1
  do 20 j=i+1,n
    A(i,j)=A(j,i)
20  continue
if(iprt .ne. 0)then
  write(nout,*)' Matrix WITH  Corr C from sfload'
  call sprtmat(A,n,n,np)
endif
call TRED2(A,n,NP,D,E)
if(iprt .ne. 0)then
  write(nout,*)' Results from tred2: tridiagonal matrix'
  call sprtmat(A,n,n,np)
endif
call TQLI(D,E,N,NP,A)
c----Check for negative eigenvalues and if small set them to zero
do 50 i=1,n
  if(D(i) .lt. 0.0)then
    if( abs(D(i)) .le. zerolim)then
      D(i)=0.0
    else
      write(nout,*)' eigenvalue ',i,' = ',d(i)
      write(nout,*)' ***value was set to 0***'
      d(i)=0.0
    endif
  endif
50  continue
C---Sort eigenvalues in ascending order
CALL EIGSRT(D,A,N,NP)
if(iprt .ne. 0)then
  write(nout,*)' Matrix of eigenvalues of C'
cccccc  print *,(d(i),i=1,n)
  call sprtmat(D,N,1,np)
  write(nout,*)' Matrix of eigenvectors of C'
  call sprtmat(A,n,n,np)
endif
c---compute component loadings
do 100 i=1,n
  do 100 j=1,n
    aux=d(j)
    a(i,j)=a(i,j)*dsqrt(aux)
100  continue
if(iprt .ne. 0)then
  write(nout,*)' Matrix of Component loadings of C'
  call sprtmat(A,n,n,np)
endif
return
END

SUBROUTINE TRED2(A,N,NP,D,E)
C *****
C * SUBROUTINE: TRED2 *
C * *
C * REMARKS : Reduction of a symmetric matrix to a *
C * Tridiagonal Form *
C * *
C * VARIABLES *
C * ===== *
C * A : INITIAL SYMMETRIC MATRIX TO BE REDUCED*
C * FOR OUTPUT A IS REPLACED BY TRIDIAG. *
C * N : DIMENSION OF MATRIX A *
C * NP : PHYSICAL DIMENSION OF MATRIX A *
C * D : DIAGONAL ELEMENTS OF A *
C * E : OFF DIAGONAL ELEMENTS WITH E(1)=0 *
C *****
IMPLICIT REAL*8(A-H,O-Z)

```

```

DIMENSION A(NP,NP),D(NP),E(NP)
IF(N .GT. 1)THEN
  DO 18 I=N,2,-1
    L=I-1
    H=0.0
    SCALE=0.
    IF(L .GT. 1)THEN
      DO 11 K=1,L
        SCALE=SCALE+ABS(A(I,K))
11      CONTINUE
      IF (SCALE .EQ. 0)THEN
        E(I)=A(I,L)
      ELSE
        DO 12 K=1,L
          A(I,K)=A(I,K)/SCALE
12          H=H+A(I,K)**2
          CONTINUE
          F=A(I,L)
          G=-SIGN(SQRT(H),F)
          E(I)=SCALE*G
          H=H-F*G
          A(I,L)=F-G
          F=0.
          DO 15 J=1,L
            A(J,I)=A(I,J)/H
            G=0.
            DO 13 K=1,J
              G=G+A(J,K)*A(I,K)
13            CONTINUE
              IF(L .GT. J)THEN
                DO 14 K=J+1,L
                  G=G+A(K,J)*A(I,K)
14                CONTINUE
              ENDIF
              E(J)=G/H
              F=F+E(J)*A(I,J)
15            CONTINUE
              HH=F/(H+H)
              DO 17 J=1,L
                F=A(I,J)
                G=E(J)-HH*F
                E(J)=G
                DO 16 K=1,J
                  A(J,K)=A(J,K)-F*E(K)-G*A(I,K)
16                CONTINUE
              CONTINUE
17            CONTINUE
          ENDIF
          ELSE
            E(I)=A(I,L)
          ENDIF
          D(I)=H
18        CONTINUE
      ENDIF
      D(1)=0.
      E(1)=0.
      DO 23 I=1,N
        L=I-1
        IF(D(I) .NE. 0)THEN
          DO 21 J=1,L
            G=0.
            DO 19 K=1,L
              G=G+A(I,K)*A(K,J)
19            CONTINUE
              DO 20 K=1,L
                A(K,J)=A(K,J)-G*A(K,I)
20            CONTINUE
          CONTINUE
21          CONTINUE
        ENDIF
        D(I)=A(I,I)
        A(I,I)=1
        IF(L .GE. 1)THEN
          DO 22 J=1,L
            A(I,J)=0.
            A(J,I)=0.
22          CONTINUE

```



```

14      CONTINUE
        D(L)=D(L)-P
        E(L)=G
        E(M)=0.
        GOTO 1
    ENDIF
15      CONTINUE
    ENDIF
    RETURN
    END

SUBROUTINE EIGSRT(D,V,N,NP)
*****
C      * SUBROUTINE: EIGSRT *
C      * *
C      * REMARKS : SORTS THE EIGENVALUES INTO DESCENDING *
C      * ORDER AND REARRANGES THE COLUMNS OF V *
C      * (Sorts results from Jacobi or Tqli) *
C      * VARIABLES *
C      * ===== *
C      * D : EIGENVALUES *
C      * V : EIGENVECTORS *
C      * N : DIMENSION OF ARRAYS *
C      * NP : PHYSICAL DIMENSION OF ARRAYS *
*****
    IMPLICIT REAL*8(A-H,O-Z)
    DIMENSION D(NP),V(NP,NP)
    DO 13 I=1,N-1
        K=I
        P=D(I)
        DO 11 J=I+1,N
            IF(D(J) .GT. P)then
                K=J
                P=D(J)
            ENDIF
11      CONTINUE
        IF (K .NE. I)THEN
            D(K)=D(I)
            D(I)=P
            DO 12 J=1,N
                P=V(J,I)
                V(J,I)=V(J,K)
                V(J,K)=P
            ENDIF
12      CONTINUE
        ENDIF
13      CONTINUE
    RETURN
    END

SUBROUTINE SPRTMAT(A,M,N,mlimit)
*****
C      * SUBROUTINE: SPRTMAT *
C      * *
C      * REMARKS : PRINTS A MATRIX OR VECTOR(transposed) *
C      * FOR VECTOR USE N = 1 *
C      * VARIABLES *
C      * ===== *
C      * A : ARRAY TO BE PRINTED *
C      * M : # OF LINES *
C      * N : # OF COLUMNS *
C      * MLIMIT : ARRAY DIMENSION *
C      * NOUT : OUTPUT FILE *
*****
    IMPLICIT REAL*8(A-H,O-Z)
    INTEGER N,M,mlimit
    DIMENSION A(MLIMIT,MLIMIT)
    COMMON /jvga/NOUT
    C      WRITE(NOUT,8000)
    C 8000 FORMAT(/)
    C-----PRINTS A COLUMN VECTOR(TRANPOSED)
    IF (N .EQ. 1) THEN
        WRITE(NOUT,9000)(A(I,1),I=1,M)
        GO TO 9999
    ENDIF
    C-----PRINTS A MATRIX (MxN)

```

```

DO 10 I=1,M
    WRITE(NOUT,9000)(A(I,J),J=1,N)
9000    FORMAT(11(1X,E12.5,1X))
10    CONTINUE
9999    CONTINUE
RETURN
END

SUBROUTINE DECOMPOS(A,N,PIVOT)
*****
C      *
C      * SUBROUTINE: DECOMPOS
C      *
C      * REMARKS : DECOMPOSE MATRIX A = LU USING
C      *          SCALED PARTIAL PIVOTING
C      *
C      * VARIABLES
C      * =====
C      *          A(I,J) : COEFFICIENT MATRIX
C      *          B (k) : CONSTANT VECTOR
C      *          N      : # OF COLUMNS
C      *          LIMIT : ARRAY DIMENSION
C      *          PIVOT : SAVE PIVOT OPERATIONS
C      *          D(J)  : LARGEST ELEMENT IN A ROW
C      *          ROW   : STORE INITIAL PIVOT POINT
C      *          PC    : INITIAL SCALING
C      *          CI    : CHECK VALUES FOR SCALING
C      *          ITEMP : STORE VALUE FOR EXCHANGING
C      *          TEMP  : " " " "
C      *          XM    : RATIO USED IN REDUCTION
C      *
C      *
C      *****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER PIVOT,N,ROW,I,J,K,ITEMP
      REAL*8 A,D,CI,PC,TEMP,XM
      PARAMETER (LIMIT=70)
      DIMENSION A(LIMIT,LIMIT),PIVOT(LIMIT),D(LIMIT)
      common /jvga/nout
C*****
C      DETERMINE LARGEST ELEMENT IN EACH ROW d(i)
C*****
c      call sprtmat(a,n,n)
      DO 10 I=1,N
          PIVOT(I)=I
          D(I)=0.0
          DO 10 J=1,N
c              write(nout,*)' dabs a(1,j)',DABS(A(I,J)),' d(i)',D(I)
              IF (DABS(A(I,J)) .GT. DABS(D(I))) D(I)=A(I,J)
          10 CONTINUE
c          do 20 i=1,n
c              if (d(i) .eq. 0.0)then
c                  write (*,*)' **maximum equal zero in row= ',i
c              endif
c          20 continue
C*****
C      START OF REDUCTION LOOP IN MATRIX DECOMPOSITION
C*****
      DO 1000 K=1,N-1
C----- START OF ALGORITHM FOR PIVOTING
          PC=DABS(A(K,K)/D(K))
          ROW=K
C----- IDENTIFY MAX. VALUE FOR REMAINING ROWS
          DO 200 I=K+1,N
              CI=DABS(A(I,K)/D(I))
              IF (CI .GT. PC) THEN
                  PC=CI
                  ROW=I
              ENDIF
          200 CONTINUE
C----- SWITCH ELEMENTS FOR NEW PIVOT ROW
          IF (ROW .GT. K) THEN
              ITEMP=PIVOT(K)
              PIVOT(K)=PIVOT(ROW)
              PIVOT(ROW)=ITEMP
              TEMP=D(K)
              D(K)=D(ROW)

```



```

        D(ROW)=TEMP
        DO 300 J=1,N
            TEMP=A(K,J)
            A(K,J)=A(ROW,J)
            A(ROW,J)=TEMP
300      CONTINUE
        ENDIF
C*****
C      SOLVE FOR UPPER & LOWER MATRIX DECOMPOSITION
C*****
        DO 400 I=K+1,N
            XM=A(I,K)/A(K,K)
            A(I,K)=XM
            DO 400 J=K+1,N
                A(I,J)=A(I,J)-XM*A(K,J)
400      CONTINUE
1000   CONTINUE
CCC    write(*,*)' Normal end Subroutine Decompos'
        RETURN
        END

SUBROUTINE SAUGMAT(A,NR,NC)
C      *****
C      * SUBROUTINE: SAUGMAT *
C      * *
C      * REMARKS : THIS SUBROUTINE AUGMENTS A MATRIX OF *
C      * OBSERVATIONS OF DEPENDENT VARIABLES TO *
C      * INCLUDE A INTERCEPT COEFFICIENT bo OF *
C      * THE REGRESSION EQUATION *
C      * *
C      * VARIABLES *
C      * ===== *
C      * A(NR,NC) : MATRIX A *
C      * NR : # OF ROWS OF A *
C      * NC : # OF COLUMNS OF A *
C      *****
        IMPLICIT REAL*8(A-H,O-Z)
        parameter (limit=70)
        common /jvga/nout
        DIMENSION A(limit,limit)
        DO 10 I=1,NR
            DO 20 J=NC+1,2,-1
                A(I,J)=A(I,J-1)
20      CONTINUE
10     A(I,1)=1.0
        CONTINUE
CCC    WRITE(*,*)' <NORMAL END SUBROUTINE "Saugmat"> '
        RETURN
        END

SUBROUTINE SINVMAT(N,A,AINV)
C      *****
C      * SUBROUTINE: SINVMAT *
C      * *
C      * REMARKS : THIS SUBROUTINE INVERTS A MATRIX *
C      * VARIABLES *
C      * ===== *
C      * A( N, N) : MATRIX A *
C      * N : ORDER OF A *
C      * PIVOT(N) : *
C      * X(N) : *
C      * B(N) : *
C      * AINV(N) : INVERSE OF A *
C      *****
        IMPLICIT REAL*8(A-H,O-Z)
        integer pivot
        parameter (limit=70)
        common /jvga/nout
        DIMENSION A(limit,limit),AINV(limit,limit),PIVOT(limit),X(LIMIT),
        *B(LIMIT)
C
C-----Decompose matrix A into upper and lower components
        CALL DECOMPOS(A,N,PIVOT)

```

```

C---- Set initial elements in identity vector
      DO 10 I=1,N
10    B(I)=0.0

C---- Loop for each column of inverse matrix
      DO 100 J=1,N
C-----Set only nonzero value
          B(J)=1.0
          CALL SOLVE(A,N,PIVOT,B,X)
          B(J)=0.0
          DO 100 I=1,N
              AINV(I,J)=X(I)
100    CONTINUE
CCC    WRITE(*,*)' <NORMAL END SUBROUTINE "sinvmat"> '
      RETURN
      END

SUBROUTINE SMULTMAT(A,MA,NA,B,MB,NB,C)
*****
C      * SUBROUTINE: SMULTMAT *
C      * * * * *
C      * REMARKS : THIS SUBROUTINE PERFORMS MATRIX MULTIPLI*
C      *          CATION: Ax=B *
C      * VARIABLES *
C      * ===== *
C      * A(MA,NA) : MATRIX A *
C      * MA : # OF ROWS OF A *
C      * NA : # OF COLUMNS OF A *
C      * B(MB,NB) : MATRIX B *
C      * MB : # OF ROWS OF B *
C      * NB : # OF COLUMNS OF B *
C      * C(MA,NB) : PRODUCT MATRIX C *
C      * * * * *
      IMPLICIT REAL*8(A-H,O-Z)
      parameter (limit=70)
      common /jvga/nout
      DIMENSION A(limit,limit),B(limit,limit),C(limit,limit)
C
C---- Check compatibility of matrices
      IF (NA .NE. MB) THEN
          WRITE (NOUT,1)MA,NA,MB,NB
1      FORMAT (3X,' *** MATRICES ARE NOT COMPATIBLE FOR MULTIPLICATION',
* ' A(',13,'x',13,')', ' B(',13,'x',13,') ***')
          GO TO 99
      ENDIF
C
C---- Performs multiplication
      DO 10 I=1,MA
          DO 10 J=1,NB
              C(I,J)=0.0
              DO 10 K=1,NA
                  C(I,J)=C(I,J)+A(I,K)*B(K,J)
10      CONTINUE
99      CONTINUE
CCC    WRITE(*,*)' <NORMAL END SUBROUTINE "smultmat"> '
      RETURN
      END

SUBROUTINE SOLVE(A,N,PIVOT,B,X)
*****
C      * SUBROUTINE: SOLVE *
C      * * * * *
C      * REMARKS : SOLVES A SYSTEM OF THE FORM LU x= b *
C      *          1st step: L z = b *
C      *          2nd step: U x = z *
C      * * * * *
C      * VARIABLES *
C      * ===== *
C      * A(I,J) : COEFFICIENT MATRIX *
C      * B (J) : CONSTANT VECTOR *
C      * N : # OF COLUMNS *
C      * LIMIT : ARRAY DIMENSION *

```

```

C      *      ROW  : STORE INITIAL PIVOT POINT      *
C      *      SUM  : LINE SUMMATION                *
C      *      z(J) : TEMPORARY SOLUTION = U x      *
C      *      x(J) : SOLUTION VECTOR              *
C      *
C      *****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER PIVOT,ROW
      PARAMETER (LIMIT=70)
      DIMENSION A(LIMIT,LIMIT),B(LIMIT),X(LIMIT),Z(LIMIT),PIVOT(LIMIT)
C*****
C      LOOP TO DETERMINE TEMPORARY VECTOR Z BY FORWARD SUBSTITUTION
C*****
      ROW=PIVOT(1)
      Z(1)=B(ROW)
      DO 200 K=2,N
          ROW=PIVOT(K)
          SUM=B(ROW)
          DO 100 J=1,K-1
              SUM=SUM-A(K,J)*Z(J)
          100 CONTINUE
          Z(K)=SUM
      200 CONTINUE
C*****
C      SOLVE FOR X USING BACK-SUBSTITUTION
C*****
      X(N)=Z(N)/A(N,N)
      DO 400 K=N-1,1,-1
          SUM=Z(K)
          DO 300 J=K+1,N
              SUM=SUM-A(K,J)*X(J)
          300 CONTINUE
          X(K)=SUM/A(K,K)
      400 CONTINUE
CCC      write(*,*)'Normal end Solve'
      RETURN
      END

      SUBROUTINE STRANMAT(A,NR,NC,C)
C      *****
C      *      SUBROUTINE: STRANMAT                    *
C      *
C      *      REMARKS   : THIS SUBROUTINE COMPUTES THE TRANSPOSE *
C      *                  OF THE MATRIX A                *
C      *      VARIABLES *
C      *      ===== *
C      *      A(NR,NC) : MATRIX A                        *
C      *      NR      : # OF ROWS OF A                  *
C      *      NC      : # OF COLUMNS OF A              *
C      *      C(NC,NR) : MATRIX C = A'                  *
C      *
C      *****
      IMPLICIT REAL*8(A-H,O-Z)
      parameter (limit=70)
      common /jvga/nout
      DIMENSION A(limit,limit),C(limit,limit)
      DO 10 I=1,NC
          DO 10 J=1,NR
              C(I,J)=A(J,I)
          10 CONTINUE
      CONTINUE
CCC      WRITE(*,*)' <NORMAL END SUBROUTINE "stranmat"> '
      RETURN
      END

      FUNCTION GASDEV(IDUM)
C      *****
C      GENERATES A GAUSSIAN DEVIATE
C      *****
      IMPLICIT REAL*8(A-H,O-Z)
      real*4 ranval
      common /jvga/nout, nouh,iseedjv
      DATA ISET/0/,INIT/0/
C-----If first time randomize initial seed for rnd
cc-----This function generate a gaussian number using F77L
c      IF (init .eq. 0)xdum=rtrand()

```

```

c      init=1
c      IF (ISET .EQ. 0) THEN
c 1      V1 = 2.*RND()-1.
c          V2 = 2.*RND()-1.
c          R = V1*V1+V2*V2
c          IF (R .GT. 1.) GOTO 1
c          FACT = SQRT(-2.*LOG(R)/R)
c          GSET=V1*FACT
c          GASDEV=V2*FACT
c          ISET=1
c      ELSE
c          GASDEV=GSET
c          ISET=0
c      ENDIF
c---This works for MS-Fortran ver 5.1
c      IF (init .eq. 0)then
c          if(iseedjv .eq. 0)then
c              call gettim(ihr,imin,isec,i100th)
c          else
c              i100th=iseedjv
c          endif
c          call seed(i100th)
c          write(nout,*)' INITIAL SEED FOR GASDEV2 IS :',i100th
c          write(nout,*)' '
c          init=1
c      ENDIF

1      IF (ISET .EQ. 0) THEN
c          call random(ranval)
c          V1 = 2.*ranval-1.
c          call random(ranval)
c          V2 = 2.*ranval-1.
c          R = V1*V1+V2*V2
c          IF (R .GT. 1.) GOTO 1
c          FACT = SQRT(-2.*LOG(R)/R)
c          GSET=V1*FACT
c          GASDEV=V2*FACT
c          ISET=1
c      ELSE
c          GASDEV=GSET
c          ISET=0
c      ENDIF
c---This works for MS-Fortran ver 5
c      RETURN
c      END

SUBROUTINE SCONDR(NP,CJV,CLOWB,CUPB,NVIOCL,NVIOCU,ICOD)
*****
C      * REMARKS : THIS SUBROUTINE TESTS THE ACCEPTANCE *
C      *          OF REGRESSED C'S BASED ON THE ACCEPTABLE*
C      *          PHYSICAL BOUNDS *
C      * VARIABLES *
C      *          ===== *
C      *          CJV(NP) : VECTOR OF REGRESSED Cs *
C      *          NP : # OF PIPES *
C      *          CLOWB,CUPB : LOWER AND UPPER BOUNDS FOR C *
C      *          NVIOCL,NVIOCU : # OF VIOLATIONS FOR LOWER AND UPPER *
C      *          ICOD : 0=NO VIOLATIONS; 1=REJECT REGRESSION *
*****
IMPLICIT REAL*8(A-H,O-Z)
parameter (limit=70)
DIMENSION CJV(limit)
ICOD=0
DO 10 I=1,NP
    IF(CJV(I) .LT. CLOWB)THEN
        NVIOCL=NVIOCL+1
        GOTO 20
    ENDIF
    IF(CJV(I) .GT. CUPB)THEN
        NVIOCU=NVIOCU+1
        GOTO 20
    ENDIF
10 CONTINUE
RETURN
20 CONTINUE

```

```

ICOD=1
RETURN
END

SUBROUTINE Slumpc(X,N,NP,LVEC,XL,NPLUMP)
C *****
C * SUBROUTINE: Slumpc *
C * *
C * REMARKS : THIS SUBROUTINE AUGMENTS A MATRIX OF *
C * OBSERVATIONS OF DEPENDENT VARIABLES TO *
C * INCLUDE A INTERCEPT COEFFICIENT b0 OF *
C * THE REGRESSION EQUATION *
C * *
C * VARIABLES *
C * ===== *
C * A(NR,NC) : MATRIX A *
C * NR : # OF ROWS OF A *
C * NC : # OF COLUMNS OF A *
C *****
IMPLICIT REAL*8(A-H,O-Z)
parameter (limit=70,nump=51)
common /jvga/nout
DIMENSION X(limit,limit),XL(limit,limit),LVEC(nump)
ccc print *,' Starting slumpc n= ',n,' np= ',np
DO 5 I=1,N
DO 5 J=1,NP
XL(I,J)=0.0
5 CONTINUE
ccc print *,' End of do 5 '
ind=LVEC(1)
ccc print *,' ind =',ind
DO 10 I=1,N
JL=1
INC=0
ccc DO 10 J=1,np
print *,' Inside do 10 i,j,inc : ',i,j,inc
IF((LVEC(JL)) .EQ. J) THEN
XL(I,ind)=XL(I,ind)+X(I,J)
IF(JL .EQ. 1)INC=INC+1
JL=JL+1
ELSE
INC=INC+1
XL(I,INC)=X(I,J)
ENDIF
10 CONTINUE
NPLUMP=INC
RETURN
END

* SUBROUTINE INDATA

* SUBROUTINE GENMPNL(MPL,M,JA,JB,KN,JIJ)

* Subroutines from Lansey and Basnet(1991)

```

D3. Program to perform calibration assessment by
Monte Carlo Method.

```

C
C
C PROGRAM : JVASSMC.
C AUTHOR : JOSE ARAUJO - OSU OCT/1991
C REMARKS : COMPUTES ASSESSMENT OF CALIBRATION BY MONTE CARLO
C           SIMULATIONS
C 11/14/91: Included condition MonteCarlo (2 criteria)subr. Snoisec
C
C           Last revision Dec 17, 1991
C
C VARIABLES
C =====
C NRUNS   : Number of Monte Carlo runs
C ISEEDJV : Seed for pseudo random generator
C          0 will start from clock i100th
C RCORC   : y - generate correlated noise for C
C          n - generate uncorrelated noise
C rsavec  : y - save perturbed C's for posterior analysis
C          (file name: for075.dat)
C nmax    : define array dimension
C iopt    : define level of output from subroutine SIM
C          0 = restricted output, 1 = full output
C          (output of geometry data from INDATA can be disabled
C          by setting nsd = 1 in file for009.dat)
C fsavec  : name of file to save C's
C nsavec  : number of unit to save C's
C icondmc: 0 no restriction for noise on C
C          1 Condition MonteCarlo-2 criteria
C xnstd   : Number of std to reject noise
C cmin,cmax : Minimum and maximum bounds to rejectC
C KN      : # of nodes (Kypipe routine)
C KK      : # of pipes ( " " )
C NPJV    : # of pipes (JVASSMC.for)
C cmeanjv : mean C from estimation
C cout    : cmeanjv + noise
C          : same as cout(used for simulation)
C stdjv   : std c from estimation
C fsaveh  : name of file to save H's
C finp    : input file for JVASSMC options
C ninp    : " unit " " "
C update 12/17/91 to compute additional statistics in statjv
C
C parameter(nmax=100)
C implicit real * 8 (a-h, o-z)
C dimension cout(nmax),cmeanjv(nmax),stdjv(nmax),
C *covmatjv(nmax,nmax),cormatjv(nmax,nmax),factorjv(nmax,nmax)
C character*1 rcorc,rsavec,fsavec*20,fsaveh*20,resp,finp*20
C COMMON/SREVISE/JS(400)
C COMMON/SINDAT/JE(100),JG(100)
C COMMON /SHAR/ BM(1100), IX(2200), IP(100,13), AL(100), GFH(700),
C . AM(3800), JP(100), JX(4500)
C COMMON/SSIMIN/ JTAN(50),EMAX(50),EMIN(50),DIAM(50),NFUL(50),
C 1 LY(32),LZ(32),E(100),KPI(100),JD(100),JF(100)
C 1,MPL(4500),NA(80),NB(80),JJUN(80),JPIP(80),
C 1 NPO(50),NJO(50),L10(10),L11(10),L12(10),XGRD(10),XGRD1(10),
C 1 XGRD2(10),DDQ,IOUT,JSKIP,KJ,KTEP,MAXT,NJOUT,NMOM,NXX
C 1 ,NPOUT,NPRP,NQ,NR,NTANK,NTEP,P,SW,TPER,TPPP
C COMMON/SALL/QEXTK(50),AA(50),BB(50),CC(50),DD(50),EE(50),FF(50),
C 1 KIP(1000),JIJ(1000),KC(100),B(100),C(100),D(100),Q(100),JC(100)
C 1,R(100),S(100),V(100),ENGY(100),JJI(100),KCL0(100),JA(100),JB(100)
C 1 ,GG(50),A1,A2,A3,A4,CQ,I3,KN,NEPS,TIME,TNEXT,TPERI,TTT,UU,NJFIX

```

```

1 , IEPS, KK, CHECK, ZQ(60), NABEL, NPUMP, NCODE
COMMON/PENAL/YMIN(100), YMAX(100)
COMMON/SNAME2/M(100)
COMMON /SINREV/ BI(100), JFIX(50), TNCD, JNCG, NPCG, NNP, LABEL, FAC,
1 XPER, NQEX
COMMON /INITBK/ INIT, LASTCL
COMMON/SSIM/Y(100), YY(100), NEX(320), NIX(320),
1 KCHNG, NTRS, IOPT, IERR
common /JVGA/NOUT, ISEEDJV
INTEGER JX, IP, M, JA, JB, IX, MPL, KC, KCLO, KIP, KPI, JIJ, JJI
c----this lines were from nlcode setup
do 1 j=1, 1100
    bm(j)=0.0
    ix(j)=0
1 continue
init=1

ninp=4
write(*,*)' Input file for JVASMC ?(CON for keyboard):'
read(*,6022)finp
open(ninp,file=finp,status='old')
write(*,*)' Number of Monte Carlo runs ? '
read(ninp,*)nruns
write(*,*)' Select initial seed for random generator:'
write(*,*)' (enter 0 for arbitrary seed )'
read(ninp,*)iseedjv
5 write(*,6000)
6000 format(' Do you want to generate correlated noise for C ?(Y/N)')
read(ninp,6001)rcorc
6001 format(a1)
if((rcorc .ne. 'Y') .and. (rcorc .ne. 'y') .and. (rcorc .ne. 'N')
*.and. (rcorc .ne. 'n'))goto 5
icondmc=0
xnstd=0.0
write(*,*)' Condition Monte Carlo (Y/N) ? '
read(ninp,6001)resp
if ((resp .eq. 'y') .or. (resp .eq. 'Y'))then
    icondmc=1
    write(*,*)' Number of STD to reject noise ?'
    read(ninp,*)xnstd
endif
write(*,*)' Give minimum acceptable C '
read(ninp,*)cmin
write(*,*)' Give maximum acceptable C '
read(ninp,*)cmax

6 WRITE(*,*)' Save C perturbed for posterior statistics ?(Y/N) '
read(ninp,6001)rsavec
if((rsavec .ne. 'Y') .and. (rsavec .ne. 'y') .and. (rsavec.ne.'N')
*.and. (rsavec .ne. 'n'))goto 6
6022 format(a20)
write(*,*)' Select desired level of simulation results '
write(*,*)' (0) restricted output; (1) full output : '
read(ninp,*)iopt
c----define units for input/output files
c unit 4 : input file for jvasmc simulation options
c unit 6 : main output from jvasmc.for
c unit 9 : network geometry input file (Kypipe format)
c 10 : " simulation results (Kypipe output)
c 66 : results from parameter simulation (input file)
c 75 : save perturbed C's
c 76 : save simulated H's
n66=66
nout=6
nsavec=75
nsaveh=76
fsavec='for075.dat'
fsaveh='for076.dat'
open(unit=nout,file='for006.dat')
open(unit=10,file='for010.dat')
c---open temp files used by Kypipe modules
open(45)
open(49)
c----write simulation options
call getdat(iyr,imon,iday)

```

```

        call gettim(ihr,imin,isec,i100th)
        write(nout,7000)imon,iday,iyr,ihr,imin
7000 format(5x, 'PROGRAM JVASSMC, Perform Assessment of calibration',
*' by Monte Carlo method',/,5x,'Execution: ',i2,'/',i2,'/',i4,
*3x,i2,':',i2,/)
        write(nout,7001)nruns,iseedjv,rcorc,rsavec,iopt,cmin,cmax
7001 format(5x,'SIMULATION OPTIONS:',/,5x,'nruns ',i4,3x,'iseedjv ',
*i4,3x,'rcorc ',a1,3x,'rsavec ',a1,3x,'iopt ',i3,' cmin ',f10.3,
*3x,' cmax ',f10.3)
        if(icondmc .eq. 1)write(nout,7002)xnstd
7002 format(5x,'*Condition Monte Carlo option * xnstd ',f10.3)
        write(*,*)' Reading network geometry - INDATA'
c--- Read network geometry (Kypipe file)
        open(unit=9,file='for009.dat',status='old')
        call indata
        close(9)
        write(*,*)' Reading results from parameter estimation'
c--- Read results from parameter estimation
        open(unit=n66,file='for066.dat',status='old')
        read(n66,6010)npjv
6010 format(i5)
        read(n66,6020)(cmeanjv(j),j=1,npjv)
6020 format(10(f13.6,1x))
        read(n66,6030)(stdjv(j),j=1,npjv)
6030 format(10(f13.8,1x))
        do 105 i=1,npjv
            read(n66,6040)(covmatjv(i,j),j=1,i)
105 continue
        read(n66,6040)tracejv
6040 format(10(e13.7,1x))
        do 110 i=1,npjv
            read(n66,6040)(cormatjv(i,j),j=1,i)
110 continue
        do 120 i=1,npjv
            read(n66,6040)(factorjv(i,j),j=1,npjv)
120 continue
        read(n66,6010)npjv2
        close(n66)
c--- check integrity of file 66
        if(npjv .ne. npjv2)then
            write(*,*)' Error in file For066 from Estimation module'
            pause
            goto 9999
        endif
        if(npjv .ne. kk)then
            write(*,*)' Number of pipes different in for009 & for066'
            pause
            goto 9999
        endif
        if((rsavec .eq. 'y') .or. (rsavec .eq. 'Y'))then
            open(unit=nsavec,file=fsavec)
            write(nsavec,6050)npjv,nruns
6050 format(i5,1x,i5)
        endif
        open(unit=nsaveh,file=fsaveh)
        write(nsaveh,6050)kn,nruns
c-----MONTE CARLO LOOP
        write(*,*)' Performing Monte Carlo Simulations'
        DO 1000 ILOOP=1,NRUNS
c-- Generate noise in coefficient C
        call snoisec(npjv,nmax,stdjv,factorjv,rcorc,cmeanjv,cout,
* icondmc,xnstd,cmin,cmax)
c---- Update C for simulation
        do 200 i=1,npjv
            c(i)=cout(i)
200 continue
c--- Save C if desired
        if((rsavec .eq. 'y') .or. (rsavec .eq. 'Y'))then
            write(nsavec,6060)(cout(i),i=1,npjv)
6060 format(10(f13.6,1x))
        endif
c---- Perform network hydraulic simulation
c----- Resetting nrts to allow more than 8 executions of SIM
        ntrs=0
ccc write(*,*)' Calling Sim'

```



```

      call sim
ccc   write(*,*)' Saving Hs'
c---- Save H's
      write(nsavch,6070)(y(i),i=1,kn)
6070  format(10(e13.8,1x))
1000  CONTINUE
      close(nsavch)
      close(nsavh)
c-----Perform statistics on H's
ccc   nnjv=kn
      write(*,*)' Calling statjv '
      call statjv(kn,nruns,nsavch,fsavch)
c-----If condition MC print statistics of violations
      if(icondmc .eq. 1)call snoisec(npjv,nmax,stdjv,factorjv,rcorc,
      *cmeanjv,cout,9,xnstd,cmin,cmax)
      write(*,*)' **NORMAL END JVASSMC**'
9999  continue
      stop
      end

      SUBROUTINE SNOISEC(NP,NMAX,STDC,FACTOR,RCORC,CINP,COU,icondmc,
      *xnstd,cmin,cmax)
C     *****
C     * SUBROUTINE: SNOISEC *
C     * *
C     * REMARKS : GENERATES NORMALLY DISTRIBUTED NOISE *
C     * *
C     * AND PERTURB COEFFICIENT C *
C     * *
C     * VARIABLES *
C     * ===== *
C     * NP : # of pipes *
C     * NMAX : array dimension *
C     * CINP : mean value of c from par. estimation *
C     * COU : contains cout + noise *
C     * STDC : contains std of c from par. estim. *
C     * factor: Component loading of CorC " " *
C     * *
C     * sqrt(lambda(j))*a(i,j) *
C     * RCORC : Y = generate correlated noise on c *
C     * N = " uncorrelated noise *
C     * *
C     * w : gaussian noise *
C     * x : multivariate noise *
C     * iseedjv: seed for random generator *
C     * ( 0 will start from clock ) *
C     * *
C     * icondmc: 0 no restriction for noise on C *
C     * *
C     * 1 Condition MonteCarlo-2 criteria *
C     * *
C     * 9 print statistics on violations *
C     * *
C     * xnstd : Number of std to reject noise *
C     * *
C     * cmin,cmax : Minimum and maximum bounds to rejectC*
C     *****
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER NP
      PARAMETER (NMP=100)
      DIMENSION STDC(nmax),factor(nmax,nmax),cinp(nmax),COU(NMAX),
      *w(nmp),x(nmp),iviostdl(nmp),iviostdu(nmp),ivioub(nmp),
      *iviolb(nmp),vio(4)
      character*1 rcorc,vio*20
      COMMON /JVGA/NOU,ISEEDJV
      data iviostdl/nmp*0/,iviostdu/nmp*0/,ivioub/nmp*0/,
      *iviolb/nmp*0/,vio(1)'/#nstdv lower'/,vio(2)'/#nstdv upper'/,
      *vio(3)'/cmin'/,vio(4)'/cmax'/
c----This option will be used at the end of the program just to print
c----Statistics of violations
      if (icondmc .eq. 9)goto 98
c Check array dimensions
      if(nmax .gt. nmp)then
        write(*,*)' Array dimension exceeded in snoisec '
        pause
        goto 99
      endif
2 continue
c This procedure will generate gaussian noise
do 3 i=1,np
  w(i)=gasdev(i)
3 continue
c This procedure will generate correlated noise

```

```

do 1000 k=1,np
  x(k)=0.
  do 2000 j=1,np
    x(k)=x(k)+w(j)*factor(k,j)
  2000 continue
  1000 continue
C-----GENERATE NOISE MEAN 0, STD=STDC(i)
DO 10 I=1,NP
c-----check if c is within feasible limits
IF ((CINP(i) .LT. cmin) .OR. (CINP(i) .GT. cmax))GOTO 90
STD=STDC(I)
if((rcorc .eq. 'y') .or. (rcorc .eq. 'Y'))then
  error=x(i)*std
else
  error=w(i)*std
endif
Cout(i)=CINP(i)+error
  10 CONTINUE
c----Test if condition MC was selected
if(icondmc .eq. 0)goto 99
c----Test rejection criteria for all pipes
do 30 i=1,np
  31 continue
c-----Test criteria 1; noise within nstd
if (cout(i) .lt. (cinp(i)-xnstd*stdc(i))) then
  iviostdl(i)=iviostdl(i)+1
  goto 35
endif
c-----Test of criteria 2, upper and lower bounds
if (cout(i) .gt. (cinp(i)+xnstd*stdc(i))) then
  iviostdu(i)=iviostdu(i)+1
  goto 35
endif
if (cout(i) .lt. cmin)then
  iviolb(i)=iviolb(i)+1
  goto 35
endif
if (cout(i) .gt. cmax)then
  iviolb(i)=iviolb(i)+1
  goto 35
endif
c-----If no violation go to check next pipe
goto 29
c-----If violation and correlated c needs to start all over again
  35 continue
  if((rcorc .eq. 'y') .or. (rcorc .eq. 'Y'))goto 2
c-----if not correlated c's generate new noise for that pipe
  w(i)=gasdev(i)
  error=w(i)*stdc(i)
  cout(i)=cinp(i)+error
  goto 31
  29 continue
  30 continue
  GOTO 99
  90 CONTINUE
  WRITE(*,*) ' CHECK RANGE FOR C=',CINP(I),' AT PIPE ',I
  pause
  goto 99
  98 continue
c-----print statistics on violations
do 9000 i=1,np
  if(iviostdl(i) .gt. 0)write(nout,9001)vio(1),iviostdl(i),i
  if(iviostdu(i) .gt. 0)write(nout,9001)vio(2),iviostdu(i),i
  if(ivioub(i) .gt. 0)write(nout,9001)vio(3),ivioub(i),i
  if(iviolb(i) .gt. 0)write(nout,9001)vio(4),iviolb(i),i
9000 continue
9001 format(/,' Criteria ',a20,' # of violations ',i4,' pipe ',
*i4)
  99 RETURN
  END

FUNCTION GASDEV(IDUM)
(Previously listed)

```

C *****

```

C      *      Subroutine:  STATJV      *
C      *      AUTHOR   :  JOSE VICENTE GRANATO DE ARAUJO      *
C      *      DATE    :  NOV 5, 1991      *
C      *      REMARKS  :  This subroutine computes the covariance and      *
C      *                  correlation matrices of the values in a file      *
C      *                  Updated from statisti.for from April 22, 1991      *
C      *      *
C      *      VARIABLES      *
C      *      =====      *
C      *      N      :  NUMBER OF realizations      *
C      *      NP     :  NUMBER OF VARIABLES      *
C      *      X(N,NP):  MATRIX OF Values to compute statistics      *
C      *      nn    :  node number      *
C      *      q     :  node flow      *
C      *      press :  node pressure      *
C      *      *
C      *      INPUT FILE      *
C      *      =====      *
C      *      nn, q, press      *
C      *      *****      *
C      *      SUBROUTINE STATJV(kn,nruns,nsaveh,fsaveh)      *
C      *      IMPLICIT REAL*8(A-H,O-Z)      *
C      *      INTEGER LIMIT      *
C      *      PARAMETER (numv=50)      *
C      *      DIMENSION covmat(numv,numv),cormat(numv,numv),xmean(numv),      *
C      *      *hjb(numv)      *
C      *      COMMON /JVGA/NOUT,issedjb      *
C      *      character*20 fsaveh      *
ccc      write(*,*)' Entering statjv'
      open(unit=nsaveh,file=fsaveh,status='old')
      read(nsaveh,6050)np,n
6050      format(i5,1x,i5)
      if(np .ne. kn) then
          write(*,*)' Number of variables do not match in file',fsaveh
          pause
          goto 999
      endif
      if(n .ne. nruns) then
          write(*,*)' Number of runs do not match in file',fsaveh
          pause
          goto 999
      endif
c-----Compute mean of observations
c      DO 10 J=1,NP
c      10      XMEAN(J)=0.0
c      DO 20 I=1,N
c      read(nsaveh,6070)(hjb(j),j=1,np)
c 6070      format(10(e13.8,1x))
c      DO 20 J=1,NP
c      XMEAN(J)=XMEAN(J)+hjb(J)
c 20      continue
c
c      close(nsaveh)
c      write(*,*)' Opening again ',fsaveh
c      open(unit=nsaveh,file=fsaveh,status='old')
c      read(nsaveh,6050)np1,n1
c      DO 30 J=1,NP
c      30      XMEAN(J)=XMEAN(J)/N
c      write(nout,*)' Mean values '
c      write(nout,6004)(xmean(j),j=1,np)
c 6004      format(10(1x,f7.3))
c      write(nout,*)' '
c-----Compute covariance matrix
c      DO 45 I=1,Np
c      DO 45 J=1,Np
c      cormat(i,j)=0.0
c 45      COVMAT(I,J)=0.0
c      do 100 i=1,n
c      read(nsaveh,6070)(hjb(j),j=1,np)
c      DO 200 K=1,NP
c      DO 200 J=1,K
c      covmat(k,j)=covmat(k,j)+(hjb(K)-XMEAN(K))*(hjb(J)-XMEAN(J))
c 200      continue
c 100      continue
c      close(nsaveh)
c      do 300 k=1,np

```

```

c      do 300 j=1,k
c      COVMAT(K,J)=covmat(k,j)/float(n-1)
c 300  CONTINUE
c
c-----Initialize arrays
DO 45 I=1,Np
  XMEAN(i)=0.0
DO 45 J=1,Np
  cormat(i,j)=0.0
  COVMAT(I,J)=0.0
45  continue
DO 100 J=1,N
  read(nsaveh,6070)(hvj(k),k=1,np)
6070  format(10(e13.8,1x))
  DO 200 I=1,NP
    XMEAN(I)=XMEAN(I)+hvj(I)
    do 200 k=1,i
      covmat(i,k)=covmat(i,k)+hvj(i)*hvj(k)
200  continue
100  continue
close(ksaveh)
do 300 i=1,np
do 300 k=1,i
c      write(*,*)covmat(I,k)
      COVMAT(i,k)=(covmat(i,k) -xmean(i)*xmean(k)/float(n))/
&float(n-1)
300  CONTINUE
write(nout,*)' Mean values '
write(nout,6004)((xmean(j)/float(n)),j=1,np)
6004  format(10(1x,f7.3))
write(nout,*)' '
write(nout,*)'      Covariance matrix'
write(nout,*)'      ====='
call sprmat(covmat,np,np,numv)
c---- Compute trace
c      trace=0.0
c      do 48 i=1,np
c      trace=trace+covmat(i,i)
c 48  continue
c      write(nout,2000)trace
c 2000 format(/,5x,'Trace = ',f13.4,/)
c---- Compute trace---Modified 12/17/91
valmin=+1e09
valmax=-1e-09
trace=0.0
do 48 i=1,np
  trace=trace+covmat(i,i)
  if(covmat(i,i) .gt. valmax)valmax=covmat(i,i)
  if(covmat(i,i) .lt. valmin)valmin=covmat(i,i)
48  continue
traceav=trace/float(kn)
write(nout,2000)trace,traceav,valmin,valmax
2000  format(/,5x,'Trace = ',f13.4,5x,'Average Cov Hp = ',f13.4,
*5x,'Min value = ',f13.4,5x,'Max value = ',f13.4)
c--Updated 12/17 to compute additional measures
valmin=+1e09
valmax=-1e-09
trace=0.0
icont=0
do 400 i=2,np
  do 400 j=1,i-1
    icont=icont+1
    trace=trace+covmat(i,j)
    if(covmat(i,j) .gt. valmax)then
      valmax=covmat(i,j)
      imax=i
      jmax=j
    endif
    if(covmat(i,j) .lt. valmin)then
      valmin=covmat(i,j)
      imin=i
      jmin=j
    endif
400  continue
traceav=trace/float(icont)

```

```

      write(nout,2005)trace,traceav,valmin,imin,jmin,valmax,imax,jmax
2005 format(5x,'Sum lower triangular Covhp = ',f13.4,2x,
*'Average Cov Hp = ',f13.4,2x,'Min value = ',f13.4,2x,
*'at row ',i3,2x,'col ',i3,/,5x,'Max value = ',f13.4,' at row ',
*i3,2x,', col ',i3)

```

```

c-----Compute correlation matrix
      DO 49 I=1, NP
      DO 49 J=1, I
49    CORMAT(I,J)=COVMAT(I,J)/DSQRT(COVMAT(I,I)*COVMAT(J,J))
      write(nout,*) ' '
      write(nout,*) '      Correlation matrix'
      write(nout,*) '      ====='
      call sprtmat(cormat,np,np,numv)
      write(nout,*) ' '
999  continue
ccc  write(*,*)' Exiting statjv'
      RETURN
      END

```

```

SUBROUTINE SPRTMAT(A,M,N,mlimit)
  (See List in previous program)

```

* The following subroutines are from Lansey and Basnet(1991)

```

SUBROUTINE INDATA

```

```

SUBROUTINE GENMPNL(MPL,M,JA,JB,KN,JIJ)

```

```

SUBROUTINE SIM

```

```

SUBROUTINE REVISE(NSWIT, nepload)

```

```

SUBROUTINE NAME(KK,W9,J)

```

```

SUBROUTINE MA18A (A,IND,IW,N,NP,G,U,IA)

```

```

SUBROUTINE MA18B (A,IRN,IP,N,NP,AWS,AVECT,MTYPE)

```

```

SUBROUTINE MA18C(A,IRN,IP,N,NP,AGRO)

```

```

SUBROUTINE MA18D (A,IRN,IP,N,NP,AWS,NAME)

```

D4. Program to perform calibration assessment by
First Order Approximation of the variance.

```

C      PROGRAM : JVASSFO.for
C      AUTHOR  : JOSE ARAUJO - OSU NOV/1991
C      REMARKS : COMPUTES ASSESSMENT OF CALIBRATION BY FIRST ORDER
C              APPROXIMATION OF THE VARIANCE OF SIMULATED PRESSURES
C
C              Last revision Dec. 17, 1991
C
C      VARIABLES
C      =====
C      NRUNS   : Number of Monte Carlo runs
C      ISEEDJV : Seed for pseudo random generator
C              0 will start from clock i100th
C      RCORC   : y - generate correlated noise for C
C              n - generate uncorrelated noise
C      rsavec  : y - save perturbed C's for posterior analysis
C              (file name: for070.dat)
C      nmax    : define array dimension
C      iopt    : define level of output from subroutine SIM
C              0 = full output, 1 = limited output
C              (output of geometry data from INDATA can be disabled
C              as an option in the for009 input data)
C      fsavec  : name of file to save C's
C      nsavec  : number of unit to save C's
C      hjv(i,j): pressure at node J caused by using (C+deltaC) at
C              pipe I
C      dhpdc(i,j) : dHPj/dCi
C      deltaC  : finite difference used to perturb C
C
C      update 12/17/91 : Print mean H from base run (Cmean from for066)
C
C      parameter(nmax=100,nmax1=101,n50=50)
C      implicit real * 8 (a-h, o-z)
C      dimension cmeanjv(nmax),stdjv(nmax),hjv(nmax,50),
C      *covmatjv(nmax,nmax),cormatjv(nmax,nmax),factorjv(nmax,nmax),
C      *dhpdc(50,nmax)
C      dimension cmeanjv(nmax),stdjv(nmax),hjv(nmax1,n50),
C      *covmatjv(nmax,nmax),cormatjv(nmax,nmax),factorjv(nmax,nmax),
C      *dhpdc(n50,nmax)
C      character*1 rsavec,fsavec*20,fsaveh*20
C      COMMON/SREVISE/JS(400)
C      COMMON/SINDAT/JE(100),JG(100)
C      COMMON /SHAR/ BM(1100), IX(2200), IP(100,13), AL(100), GFH(700),
C      . AM(3800), JP(100), JX(4500)
C      COMMON/SSIMIN/ JTAN(50),EMAX(50),EMIN(50),DIAM(50),NFUL(50),
C      1 LY(32),LZ(32),E(100),KPI(100),JD(100),JF(100)
C      1,MPL(4500),NA(80),NB(80),JJUN(80),JPIP(80),
C      1 NPO(50),NJO(50),L10(10),L11(10),L12(10),XGRD1(10),
C      1 XGRD2(10),DDQ,IOUT,JSKIP,KJ,KTEP,MAXT,NJOUT,NMOM,NXX
C      1 ,NPOUT,NPRP,NQ,NR,NTANK,NTEP,P,SW,TPER,TPPP
C      COMMON/SALL/QEXTK(50),AA(50),BB(50),CC(50),DD(50),EE(50),FF(50),
C      1 KIP(1000),JIJ(1000),KC(100),B(100),C(100),D(100),Q(100),JC(100)
C      1,R(100),S(100),V(100),ENGY(100),JJI(100),KCLO(100),JA(100),JB(100)
C      1 ,GG(50),A1,A2,A3,A4,CQ,I3,KN,NEPS,TIME,TNEXT,TPERI,TTT,UU,NJFIX
C      1 ,IEPS,KK,CHECK,ZQ(60),NABEL,NPUMP,NCODE
C      COMMON/PENAL/YMIN(100),YMAX(100)
C      COMMON/SNAME2/M(100)
C      COMMON /SINREV/ BI(100),JFIX(50),TNCD,JNCG,NPCG,NNP,LABEL,FAC,
C      1 XPER,NQEX
C      COMMON /INITBK/ INIT,LASTCL
C      COMMON/SSIM/Y(100),YY(100),NEX(320),NIX(320),
C      1 KCHNG,NTRS,IOPT,IERR
C      common /JVGA/NOUT,ISeEDJV

```

```

      INTEGER JX,IP,M,JA,JB,IX,MPL,KC,KCLO,KIP,KPI,JIJ,JJI
c----this lines were from nlcode setup
      do 1 j=1,1100
          bm(j)=0.0
          ix(j)=0
      1   continue
          init=1

          write(*,*)' Give deltaC to perturb C ? '
          read(*,*)deltac
6001  format(a1)
      6  WRITE(*,*)' Save C perturbed for posterior statistics?(Y/N) '
          read(*,6001)rsavec
          if((rsavec .ne. 'Y') .and. (rsavec .ne. 'y') .and. (rsavec.ne.'N')
          *.and. (rsavec .ne. 'n'))goto 6
          write(*,*)' Select desired level of simulation results '
          write(*,*)'      (0) restricted output; (1) complete output : '
          read(*,*)iopt
c----define units for input/output files
c   unit 80: output from jvassfo.for
c   unit 9 : network geometry input file (Kypipe format)
c       11 : " simulation results (Kypipe output)
c       66 : results from parameter simulation (input file)
c       85 : save C's perturbed
c       86 : save H's simulated
          n66=66
          nout=6
          nsavec=85
          nsaveh=86
          fsaveh='for086.dat'
          fsavec='for085.dat'
          open(unit=nout,file='for080.dat')
          open(unit=10,file='for081.dat')
c---open temp files used by Kypipe modules
          open(45)
          open(49)
c----write input options:
          call getdat(iyr,imon,iday)
          call gettim(ihr,imin,isec,i100th)
          write(nout,7010)imon,iday,iyr,ihr,imin
      7010 format(5x,'PROGRAM JVASSFO, Perform Assessment of calibration',
          *' by First Order Aproximation of the Variance method',/,5x,
          *'Execution: ',i2,'/',i2,'/',i4,3x,i2,':',i2,/)
          write(nout,7011)deltac,rsavec,iopt
      7011 format(5x,'SIMULATION OPTIONS:',/,5x,'deltac ',f10.6,3x,
          *'rsavec ',a1,3x,'iopt ',i3)
c--- Read network geometry (Kypipe file)
          open(unit=9,file='for009.dat',status='old')
          call indata
          close(9)
c---Read results from parameter estimation
          open(unit=n66,file='for066.dat',status='old')
          read(n66,6010)npjv
      6010 format(i5)
          read(n66,6020)(cmeanjv(j),j=1,npjv)
      6020 format(10(f13.6,1x))
          read(n66,6030)(stdjv(j),j=1,npjv)
      6030 format(10(f13.8,1x))
          do 105 i=1,npjv
              read(n66,6040)(covmatjv(i,j),j=1,i)
      105  continue
          read(n66,6040)tracejv
      6040 format(10(e13.7,1x))
          do 110 i=1,npjv
              read(n66,6040)(cormatjv(i,j),j=1,i)
      110  continue
          do 120 i=1,npjv
              read(n66,6040)(factorjv(i,j),j=1,npjv)
      120  continue
          read(n66,6010)npjv2
          close(n66)

c---check integrity of file 66
          if(npjv .ne. npjv2)then
              write(*,*)' Error in file For066 from Estimation module'

```

```

        pause
        goto 9999
    endif
    if(npjv .ne. kn)then
        write(*,*)' Number of pipes different in for009 & for066'
        pause
        goto 9999
    endif
    if((rsavec .eq. 'y') .or. (rsavec .eq. 'Y'))then
        open(unit=nsavec,file=fsavec)
        write(nsavec,6050)npjv,npjv+1
6050     format(i5,1x,i5)
    endif

        open(unit=nsaveh,file=fsaveh)
        write(nsaveh,6050)kn,npjv+1
c---PHASE I - Compute DHPDC by finite differences
c-----Loop to perturb each C by deltaC
        write(*,*)' Performing Hydraulic simulations'
        do 1000 iloop=1,npjv
c----     Update C for simulation
            do 200 i=1,npjv
                c(i)=cmeanjv(i)
                200     continue
c----     Perturb C(iloop) by deltaC
                c(iloop)=c(iloop)+deltac
c---     Save C if desired
                if((rsavec .eq. 'y') .or. (rsavec .eq. 'Y'))then
                    write(nsavec,6060)(c(i),i=1,npjv)
                    6060     format(10(f13.6,1x))
                endif
c----     Perform network hydraulic simulation
                ntrs=0
c---     write(*,*)' Calling Sim'
                call sim
c---     write(*,*)' Saving Hs'
                250     Save H's
                    do 250 i=1,kn
                        hjv(iloop,i)=y(i)
                    250     continue
                    write(nsaveh,6070)(y(i),i=1,kn)
                    6070     format(10(e13.8,1x))
                    1000     continue
c-----     Perform the base run
                    do 300 i=1,npjv
                        c(i)=cmeanjv(i)
                    300     continue
                    if((rsavec .eq. 'y') .or. (rsavec .eq. 'Y'))then
                        write(nsavec,6060)(c(i),i=1,npjv)
                    endif
                    write(*,*)' Calling Sim for base run'
                    call sim
c---     write(*,*)' Saving Hs'
                    do 350 i=1,kn
                        hjv(npjv+1,i)=y(i)
                    350     continue
                    write(nsaveh,6070)(y(i),i=1,kn)
                    close(nsavec)
                    close(nsaveh)
                    write(nout,*)' Mean values of H (FOSM) '
                    write(nout,6004)(y(i),i=1,kn)
                    6004     format(10(1x,f7.3))
c-----     Compute sensitivity matrix dHp/dC
                    c      **Note Line NPjv+1 is the base run**
                        do 500 i=1,npjv
                            do 500 j=1,kn
                                Dhpdc(J,I)=(Hjv(NPjv+1,J)-Hjv(I,J))/DELTAC
                            500     CONTINUE
                                WRITE(NOUT,6011)NPjv,KN,DELTAC
                                6011     FORMAT(5x,'# OF PIPES = ',I3,/,5x,'# OF NODES = ',I3,/,
                                    *5x,'DELTAC = ',F12.6,/)
                    c---PHASE II - Compute CovHp = dHP/dC * CovC * [dHP/dC]'
                    call scovhp(kn,npjv,dhpdc,covmatjv)
                    9999     continue

```



```

write(*,*)' **Normal end JVAFFO**'
stop
end

C *****
C * Subroutine: SCOVHP *
C * AUTHOR : JOSE VICENTE GRANATO DE ARAUJO *
C * DATE : NOV 19, 1991 *
C * REMARKS : This subroutine computes the covariance matrix *
C * CovHp = dHP/dC * CovC * [dHP/dC]' *
C * VARIABLES *
C * ===== *
C * KN : NUMBER OF NODES *
C * NPJV : NUMBER OF VARIABLES(pipes) *
C * DHPDC(KN, NPJV) : Partial dHP/dC *
C * COVC(NPJV, NPJV): Lower triangular of Cov.of C from P.Est. *
C * COVHP(KN, KN) : Cov. matrix of HP = ab * dhpdcT *
C * DHPDCT(NPJV, KN): dhpdc' *
C * AB(KN, NPJV) : dhpdc * covc *
C *****
SUBROUTINE Scovhp(kn,npjv,dhpdc,covc)
IMPLICIT REAL*8(A-H,O-Z)
PARAMETER (nmax=100,n50=50)
DIMENSION dhpdc(n50,nmax),covc(nmax,nmax),covhp(n50,n50),
*dhpdcT(nmax,n50),ab(n50,nmax),cormat(n50,n50)
COMMON /JVGA/NOUT,issedjv
ccc character*20 fsaveh
ccc write(*,*)' Entering scovhp'
C-----Check dimension of arrays and solvability of system
IF ( (npjv .GT. nmax) .OR. (KN .GT. n50))THEN
WRITE(*,1)npjv,nmax,kn,n50
1 format(1x,'DIMENSION OF ARRAYS EXCEEDED IN SCOVHP ',4(i5))
write(*,*)npjv,nmax,kn,n50
STOP
ENDIF
C-----Form the upper part of matrix covc
do 30 i=1,npjv
do 30 j=i+1,npjv
covc(i,j)=covc(j,i)
30 continue
C-----Compute CovHp = dHp/dC * Cov C * [dHp/dC]'
call smultmat(dhpdc, kn,npjv, covc,npjv,npjv,AB,n50,nmax,nmax,nmax)
CALL STRANMAT(dhpdc, kn,npjv, dhpdcT,n50,nmax)
CALL SMULTMAT(AB, kn,npjv, dhpdcT,npjv, kn, COVHP, n50, nmax, nmax, n50)

C-----print input/output values
write(nout,*) ' Sensitivity matrix dHp/dC'
write(nout,*) ' ====='
call sprtmat(dhpdc, kn,npjv, n50, nmax)
write(nout,*) ' '
write(nout,*) ' Cov C matrix'
write(nout,*) ' ====='
call sprtmat(covc, npjv,npjv, nmax, nmax)
write(nout,*) ' '
write(nout,*) ' Cov Hp matrix'
write(nout,*) ' ====='
call sprtmat(COVHP, kn, kn, n50, n50)
write(nout,*) ' '
C----- Compute trace---Modified 21/07/91
valmin=+1e09
valmax=-1e-09
trace=0.0
do 48 i=1, kn
trace=trace+covhp(i,i)
if(covhp(i,i) .gt. valmax)valmax=covhp(i,i)
if(covhp(i,i) .lt. valmin)valmin=covhp(i,i)
48 continue
traceav=trace/float(kn)
write(nout,2000)trace,traceav,valmin,valmax
2000 format(/,5x,'Trace = ',f13.4,5x,'Average Cov Hp = ',f13.4,
*5x,'Min value = ',f13.4,5x,'Max value = ',f13.4)
C--Updated 20/07 to compute additional measures
valmin=+1e09

```

```

    valmax=-1e-09
    trace=0
    icont=0
    do 100 i=2, kn
      do 100 j=1, i-1
        icont=icont+1
        trace=trace+covhp(i,j)
        if(covhp(i,j) .gt. valmax)then
          valmax=covhp(i,j)
          imax=i
          jmax=j
        endif
        if(covhp(i,j) .lt. valmin)then
          valmin=covhp(i,j)
          imin=i
          jmin=j
        endif
      endif
    100 continue
    traceav=trace/float(icont)
    write(nout,2005)trace,traceav,valmin,imin,jmin,valmax,imax,jmax
2005 format(5x,'Sum lower triangular CovHp = ',f13.4,2x,
  *'Average Cov Hp = ',f13.4,2x,'Min value = ',f13.4,2x,
  *'at row ',i3,2x,'col ',i3,'/',5x,'Max value = ',f13.4,' at row ',
  *i3,2x,' col ',i3)
c-----Compute correlation matrix
    DO 149 I=1, kn
      DO 149 J=1, I
149   CORMAT(I,J)=COVHP(I,J)/DSQRT(COVHP(I,I)*COVHP(J,J))
      write(nout,*) ' '
      write(nout,*) '   Correlation matrix'
      write(nout,*) '   ====='
      call sprtmat(cormat, kn, kn, n50, n50)
      write(nout,*) ' '
999 continue
ccc  write(*,*)' Exiting scovhp'
    RETURN
    END

SUBROUTINE SMULTMAT(A,MA,NA,B,MB,NB,C,max,nax,mbx,nbx)
  (Subroutine list is shown in previous program)

SUBROUTINE STRANMAT(A,NR,NC,C,max,nax)

SUBROUTINE SPRTMAT(A,M,N,mlimit,nlimit)
  (See list in previous program)

* The subroutines below are from Lansey and Basnet (1991)

SUBROUTINE INDATA

SUBROUTINE GENMPL(MPL,M,JA,JB,KN,JIJ)

SUBROUTINE SIM

SUBROUTINE REVISE(NSWIT, nepload)

SUBROUTINE NAME(KK,W9,J)

SUBROUTINE MA18A (A,IND,IW,N,NP,G,U,IA)

SUBROUTINE MA18B (A,IRN,IP,N,NP,AWS,AVECT,MTYPE)

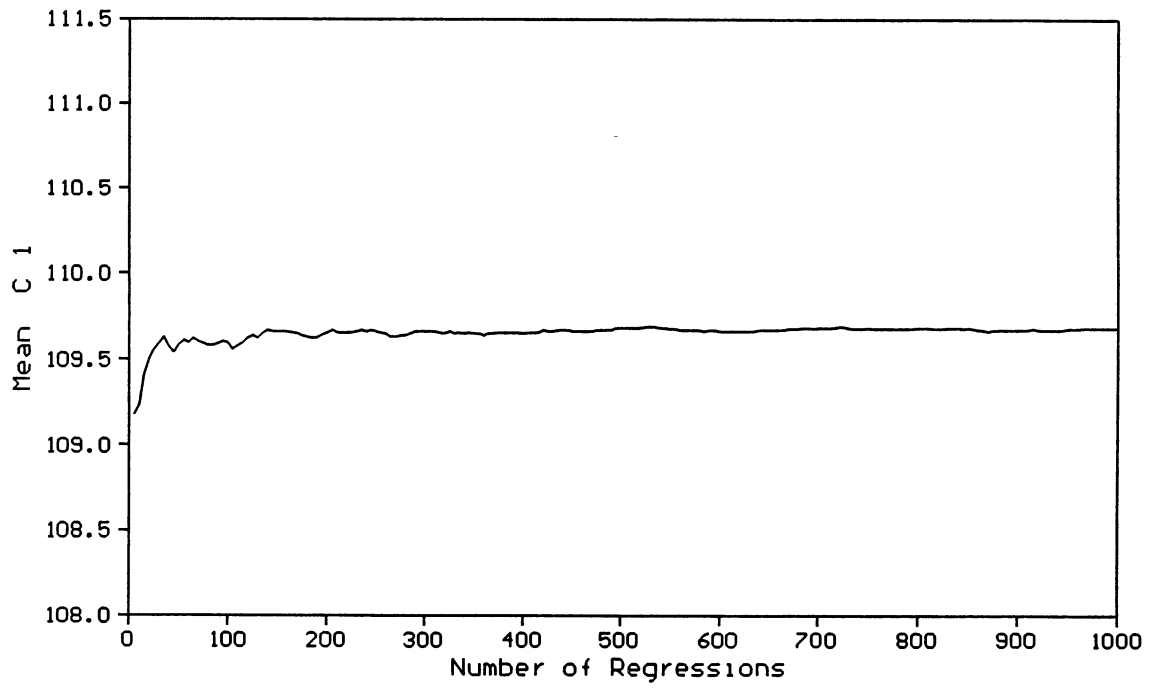
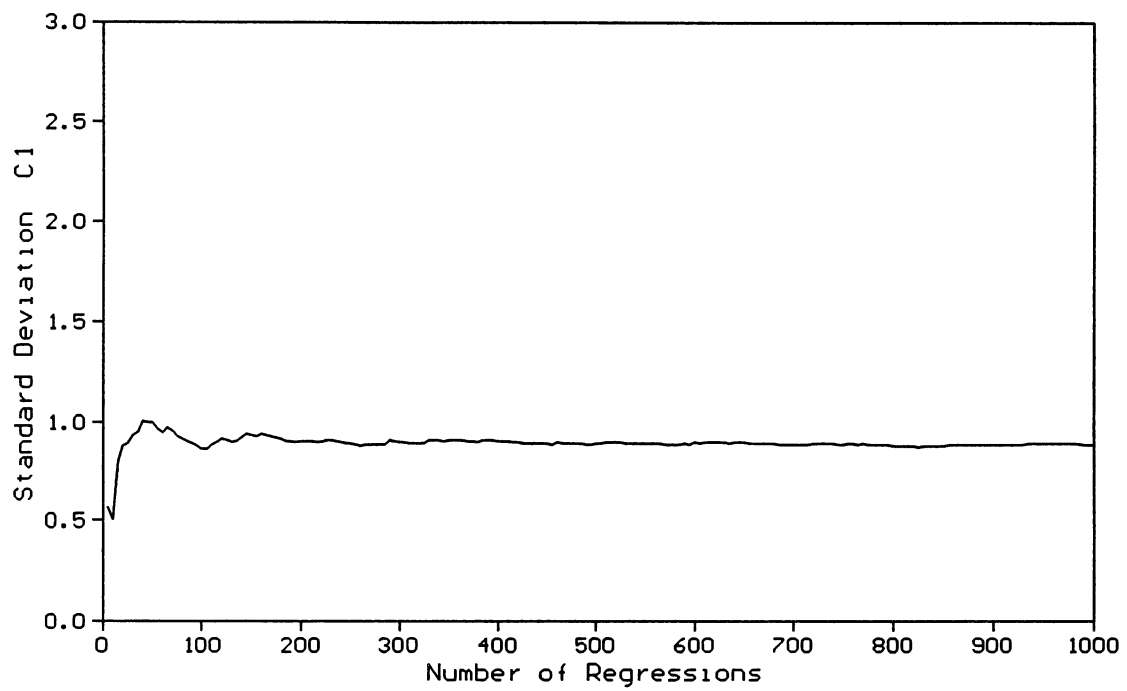
SUBROUTINE MA18C(A,IRN,IP,N,NP,AGRO)

SUBROUTINE MA18D (A,IRN,IP,N,NP,AWS,NAME)

```

APPENDIX E

STOCHASTIC CONVERGENCE OF THE MEAN AND STANDARD
DEVIATION OF ROUGHNESS PARAMETERS

Figure E.1. Stochastic Convergence of Mean C_1 Figure E.2. Stochastic Convergence of Std C_1

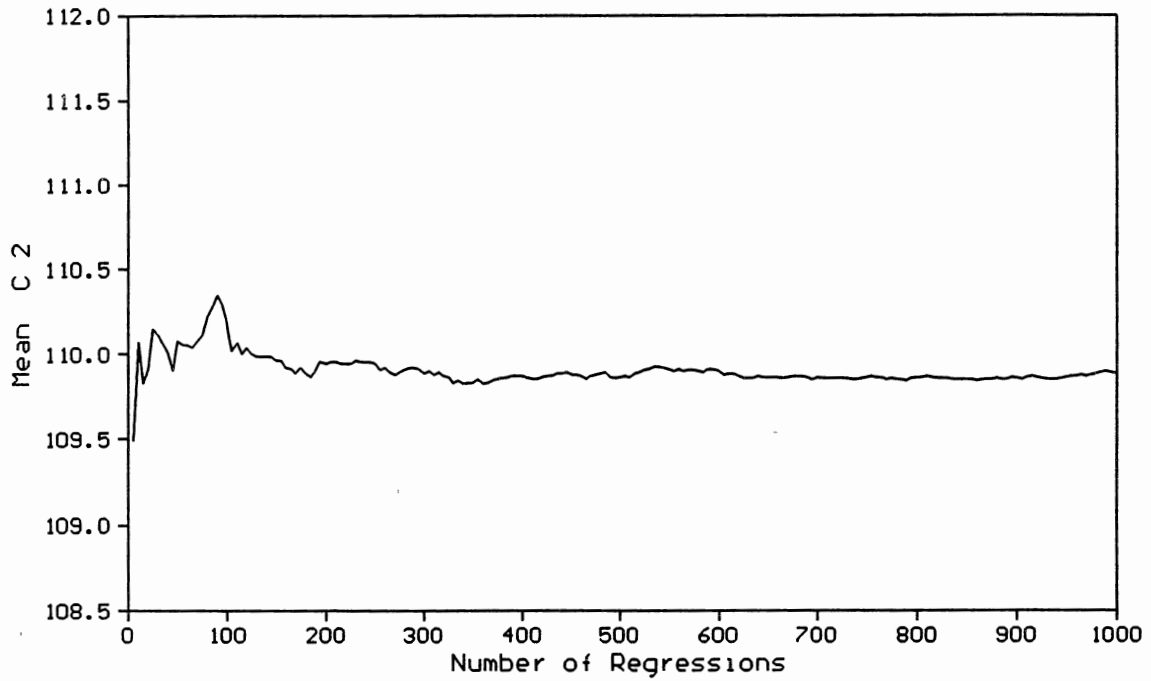


Figure E.3. Stochastic Convergence of Mean C_2

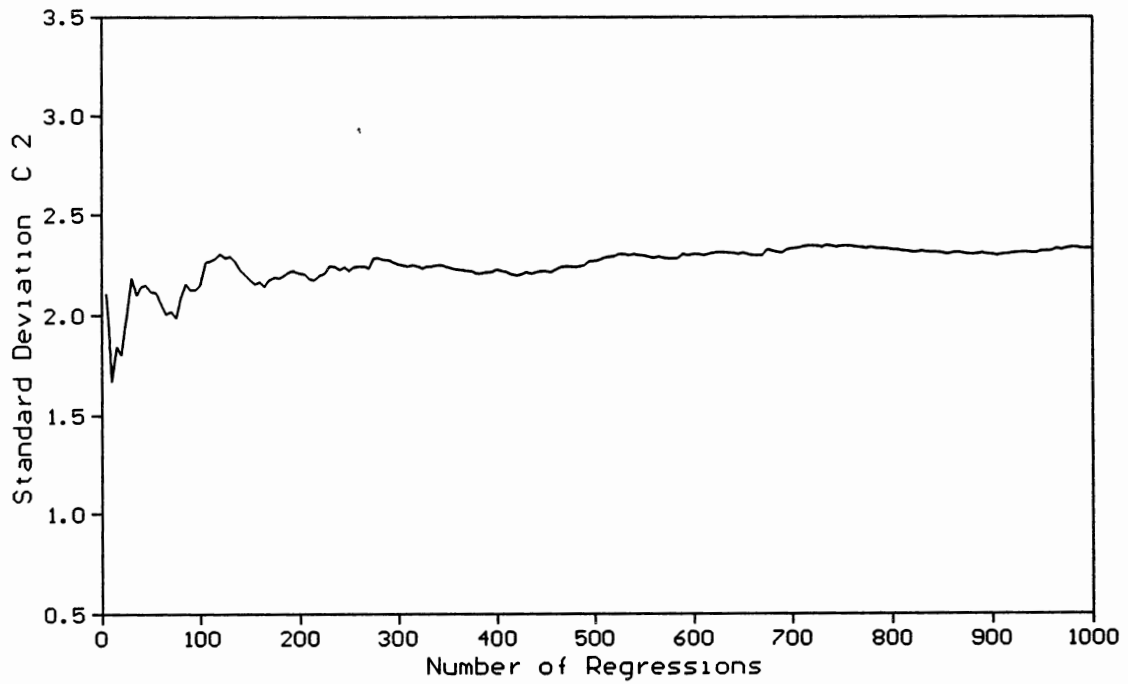
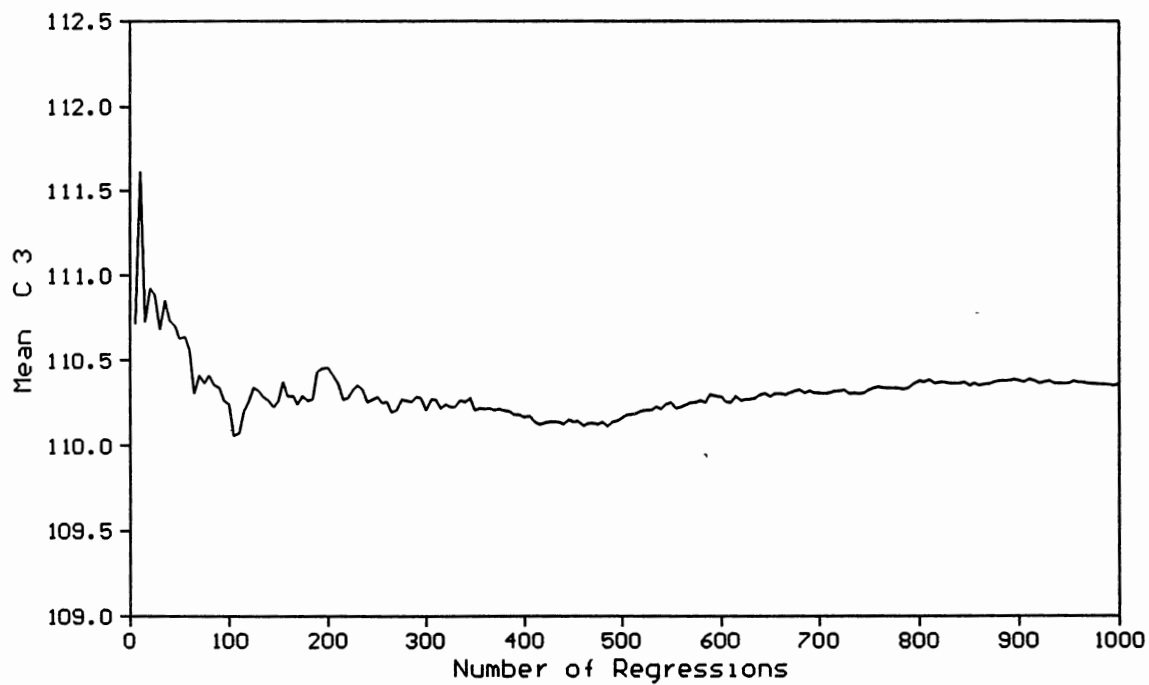
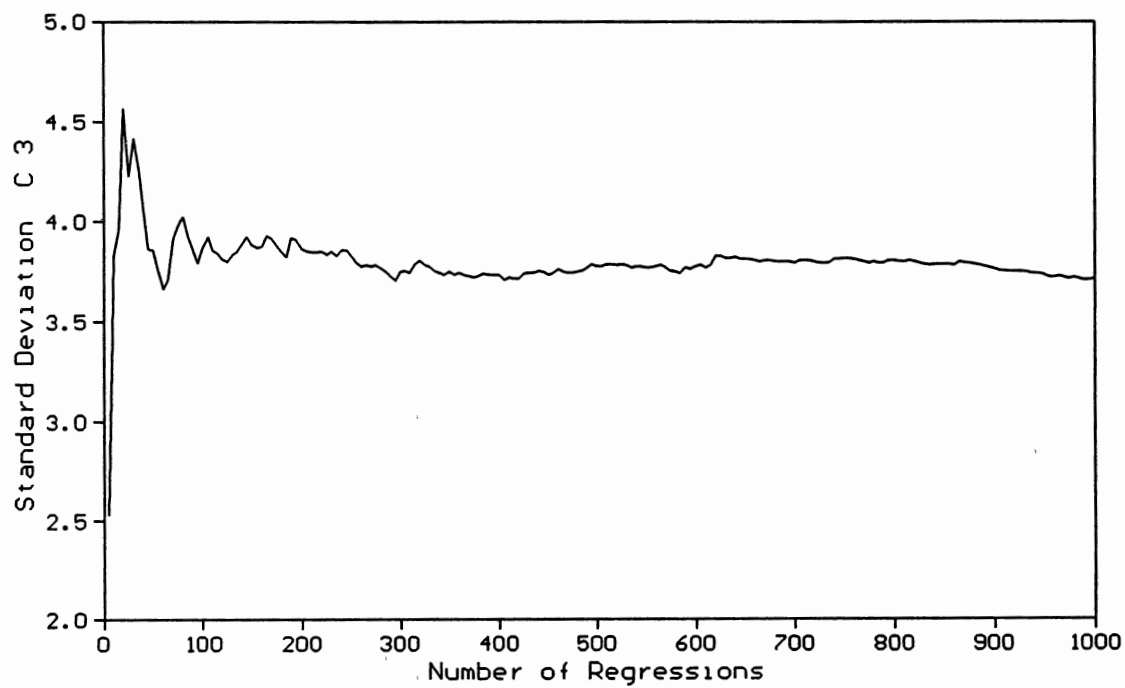
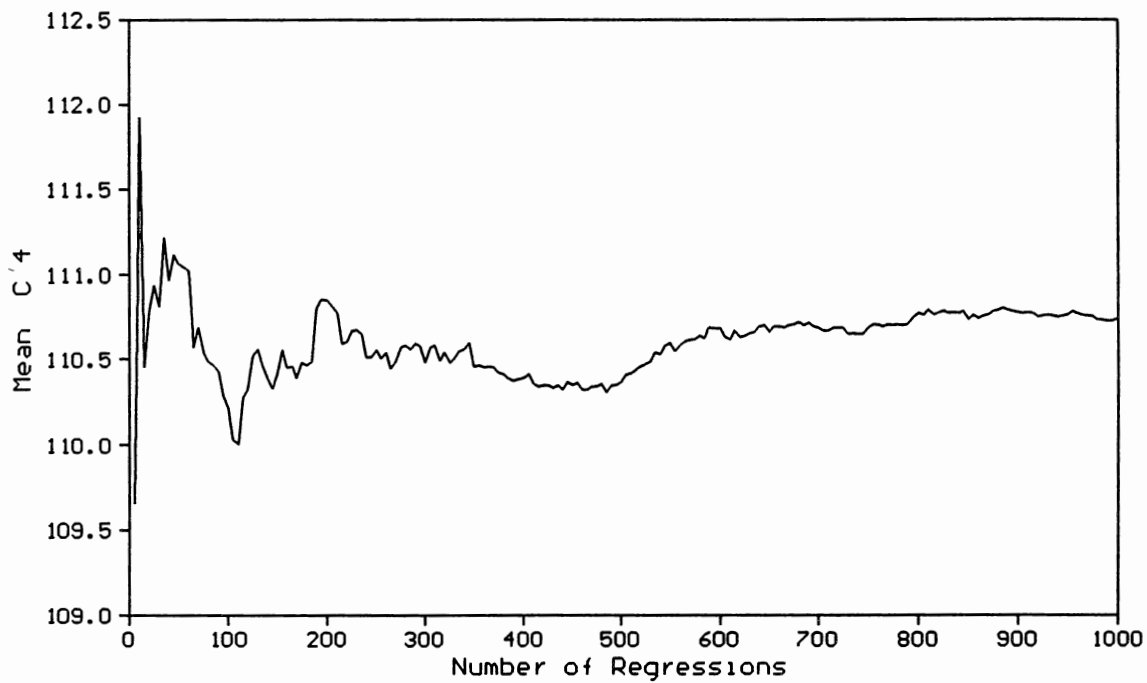
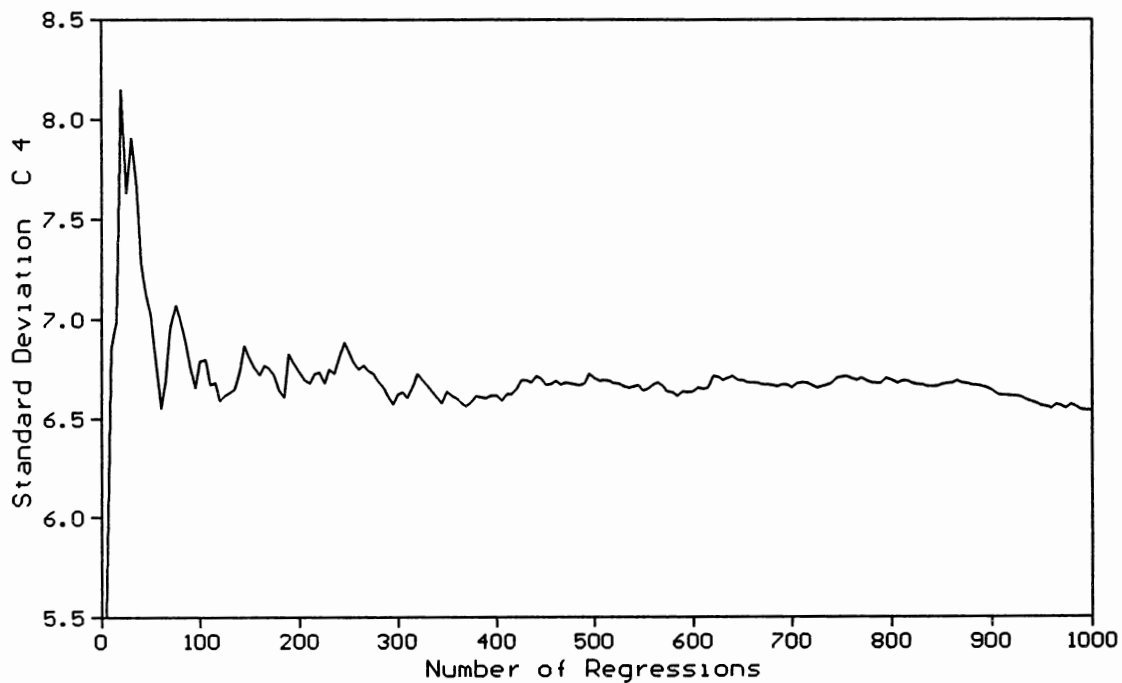


Figure E.4. Stochastic Convergence of Std C_2

Figure E.5. Stochastic Convergence of Mean C_3 Figure E.6. Stochastic Convergence of Std C_3

Figure E.7. Stochastic Convergence of Mean C_4 Figure E.8. Stochastic Convergence of Std C_4

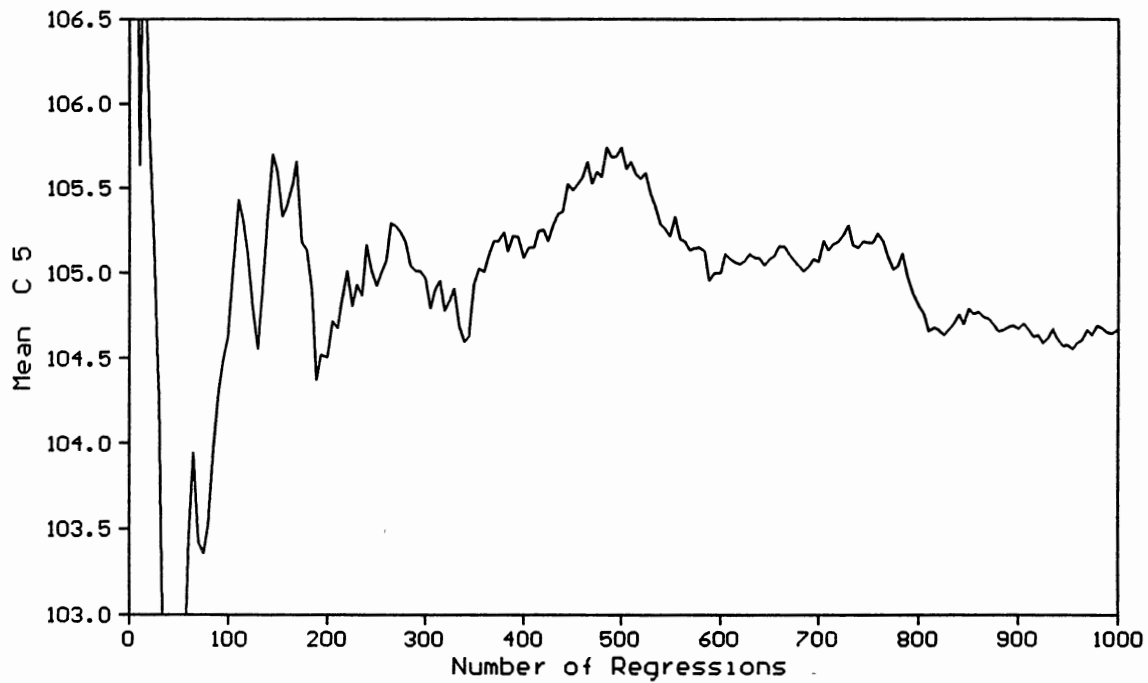


Figure E.9. Stochastic Convergence of Mean C_5

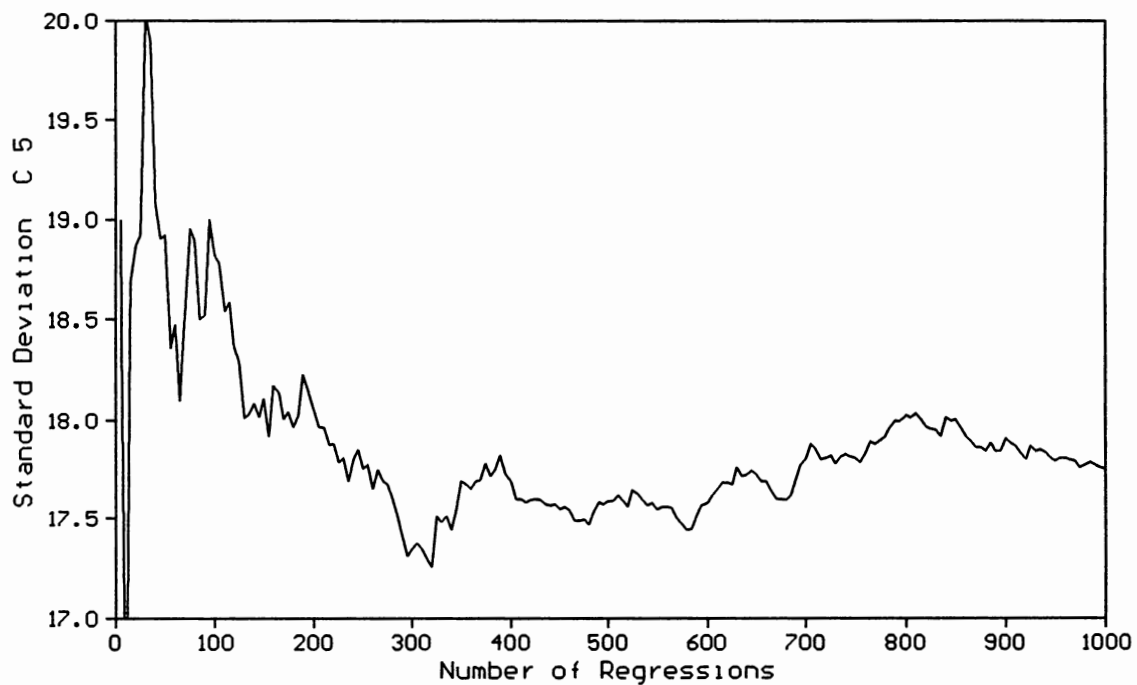


Figure E.10. Stochastic Convergence of Std C_5

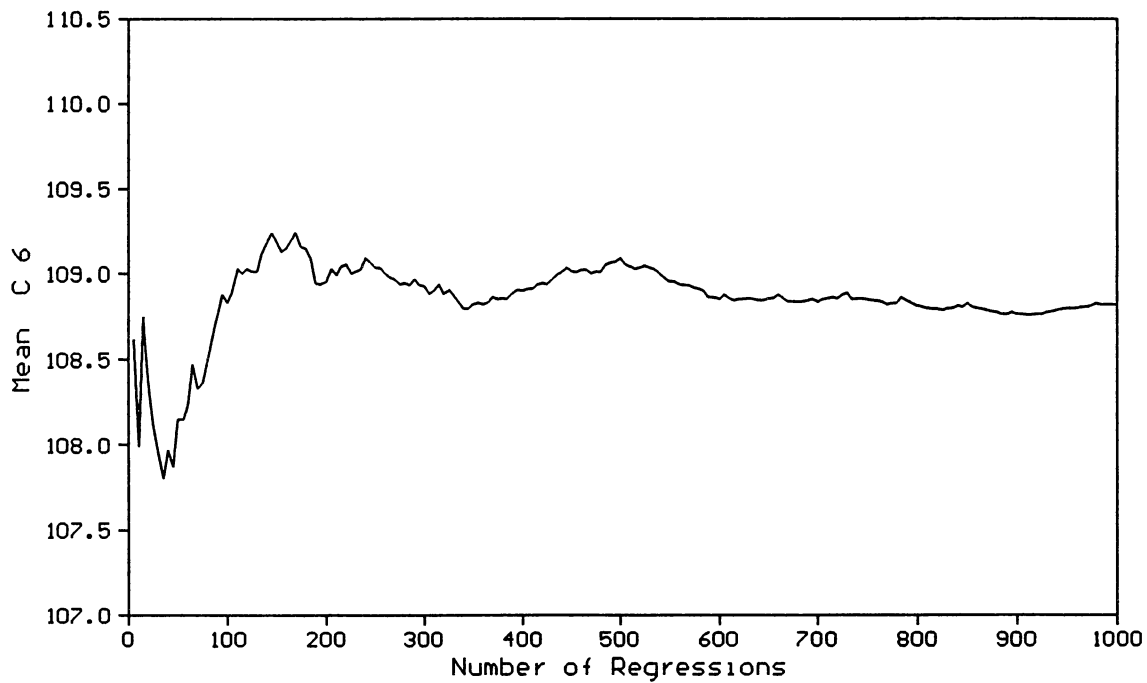


Figure E.11. Stochastic Convergence of Mean C_6

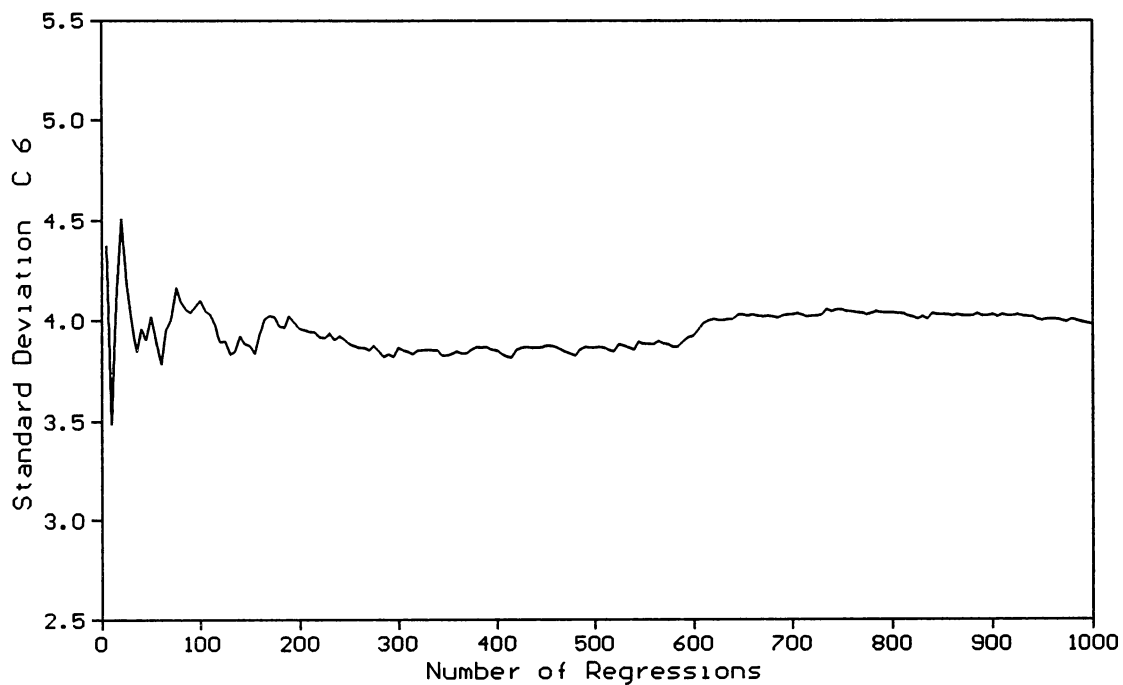


Figure E.12. Stochastic Convergence of Std C_6

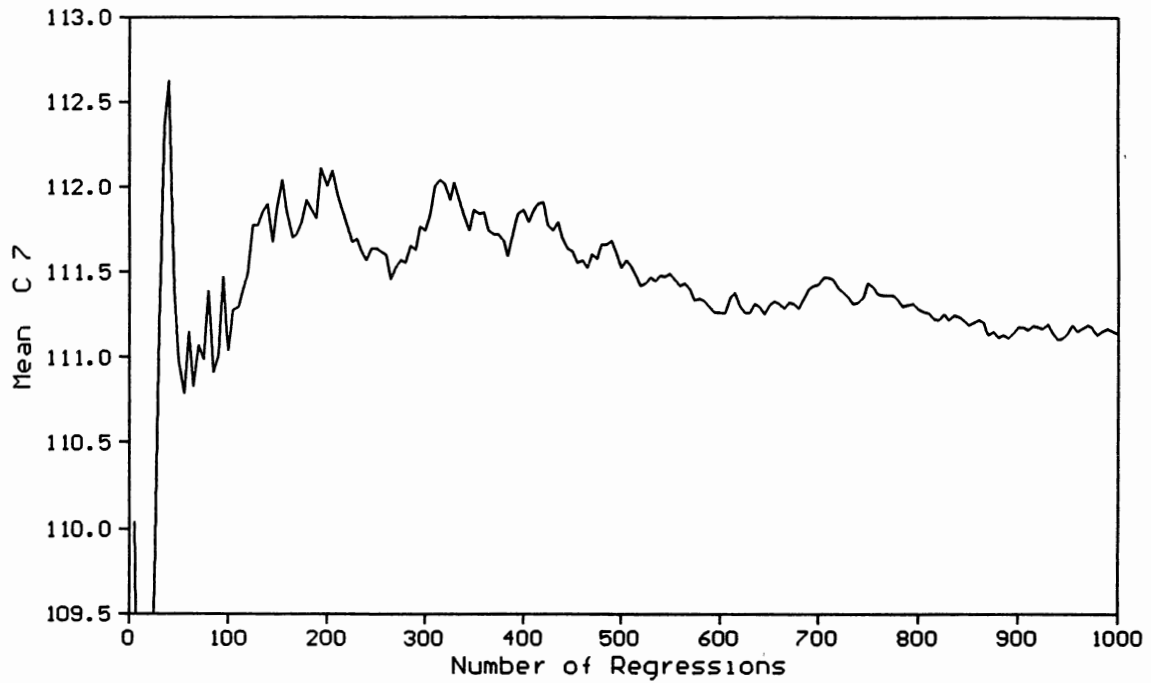
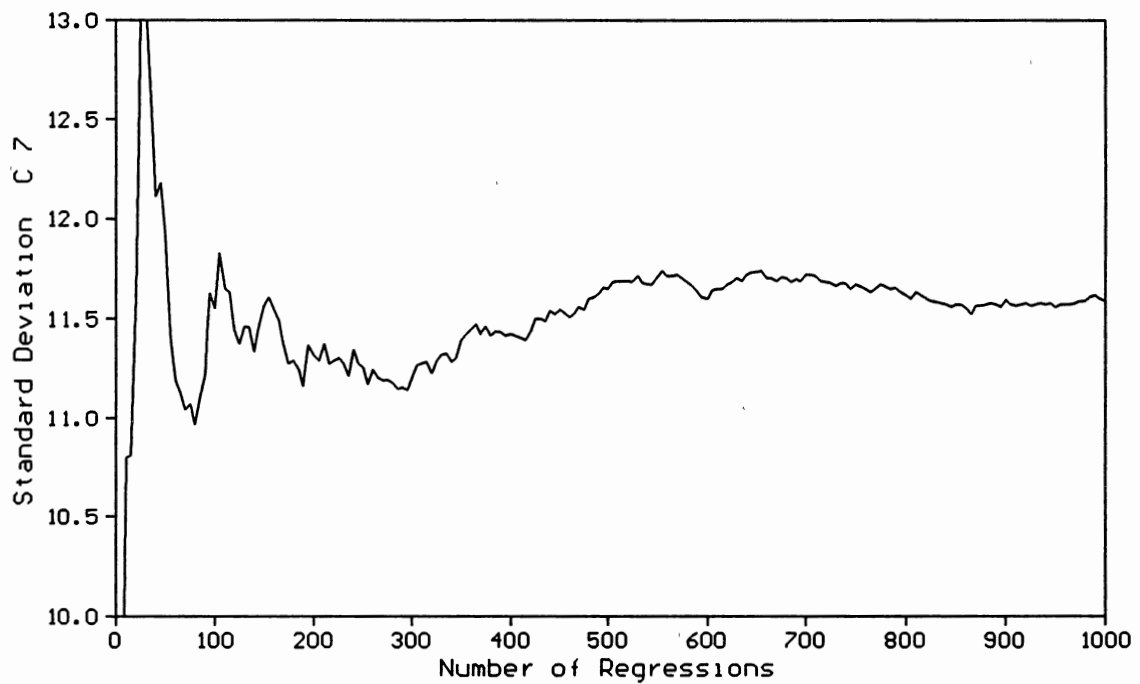
Figure E.13. Stochastic Convergence of Mean C₇Figure E.14. Stochastic Convergence of Std C₇



Figure E.15. Stochastic Convergence of Mean C₈



Figure E.16. Stochastic Convergence of Std C₈

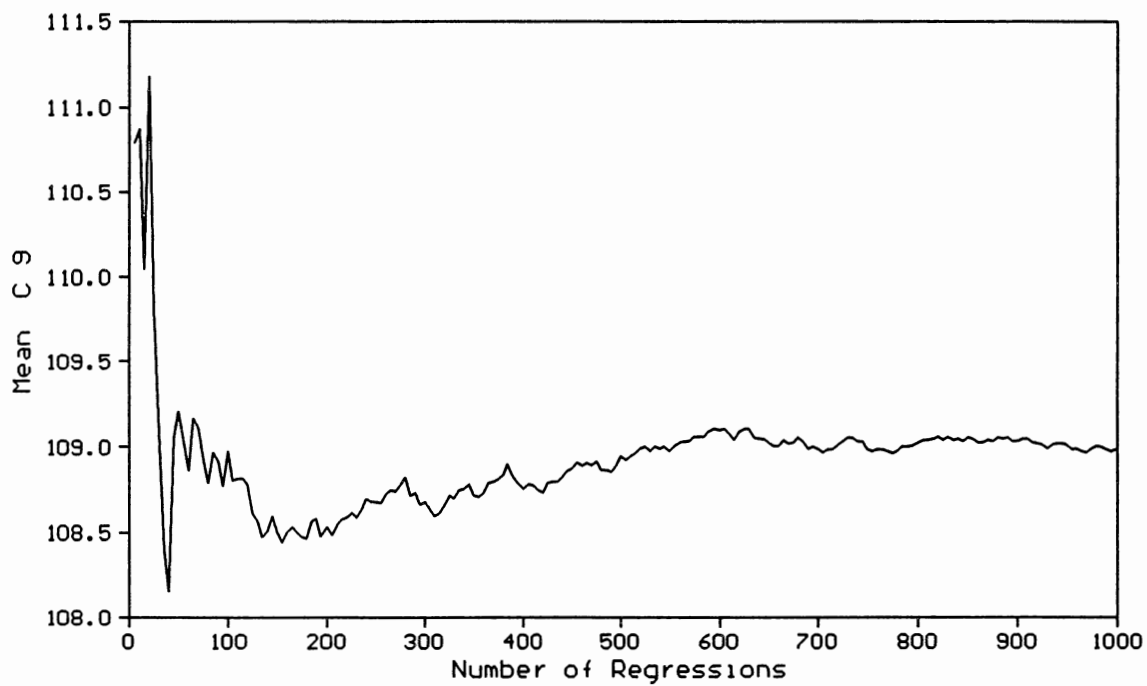


Figure E.17. Stochastic Convergence of Mean C_9

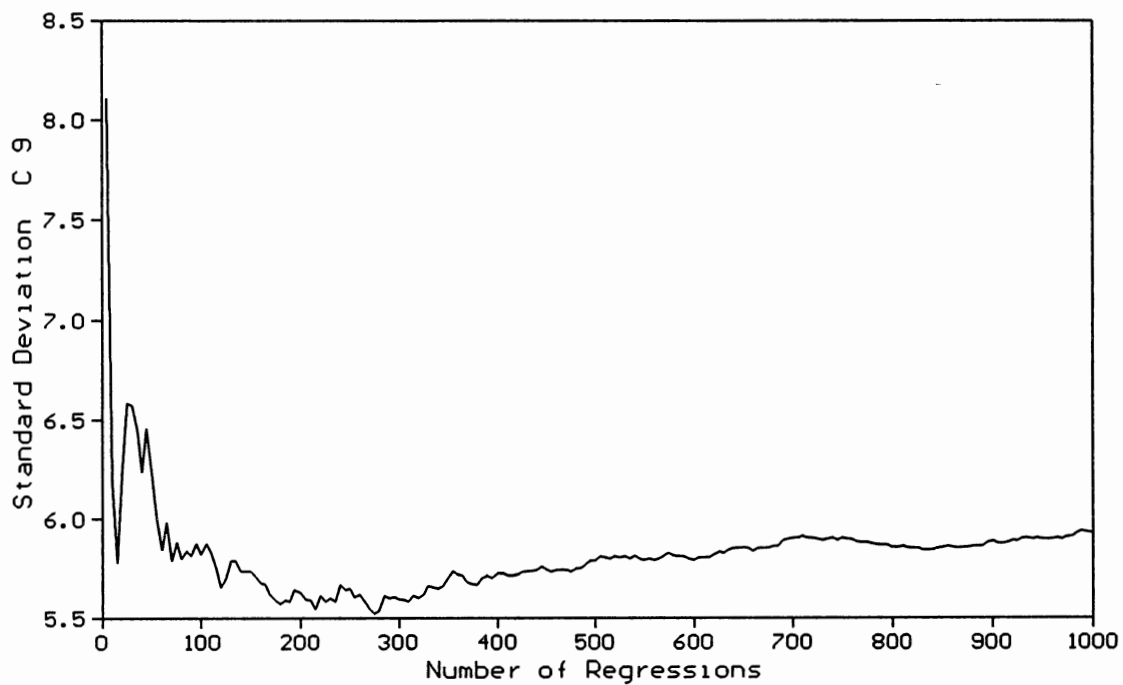


Figure E.18. Stochastic Convergence of Std C_9

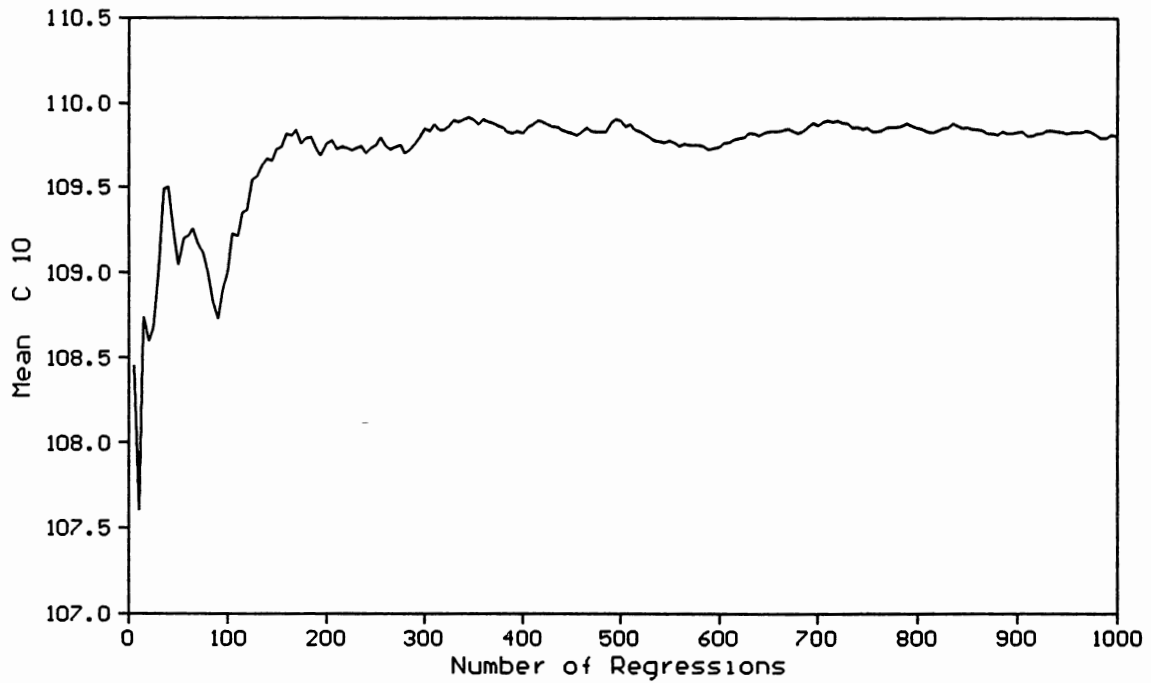


Figure E.19. Stochastic Convergence of Mean C₁₀

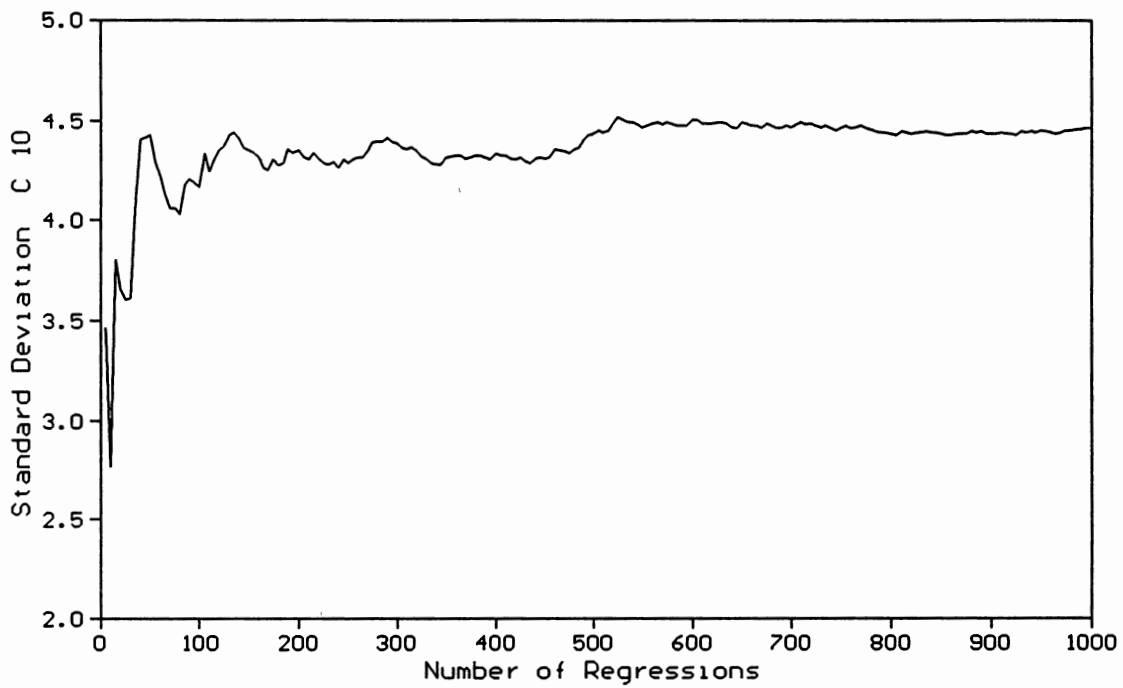


Figure E.20. Stochastic Convergence of Std C₁₀

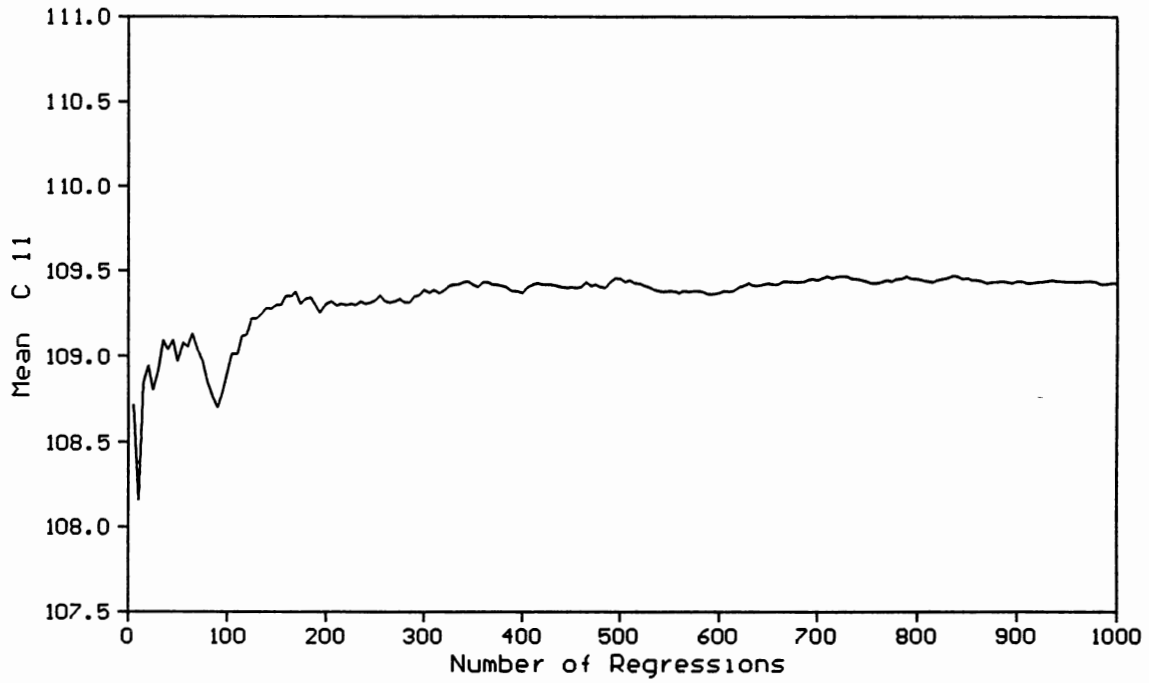


Figure E.21. Stochastic Convergence of Mean C₁₁

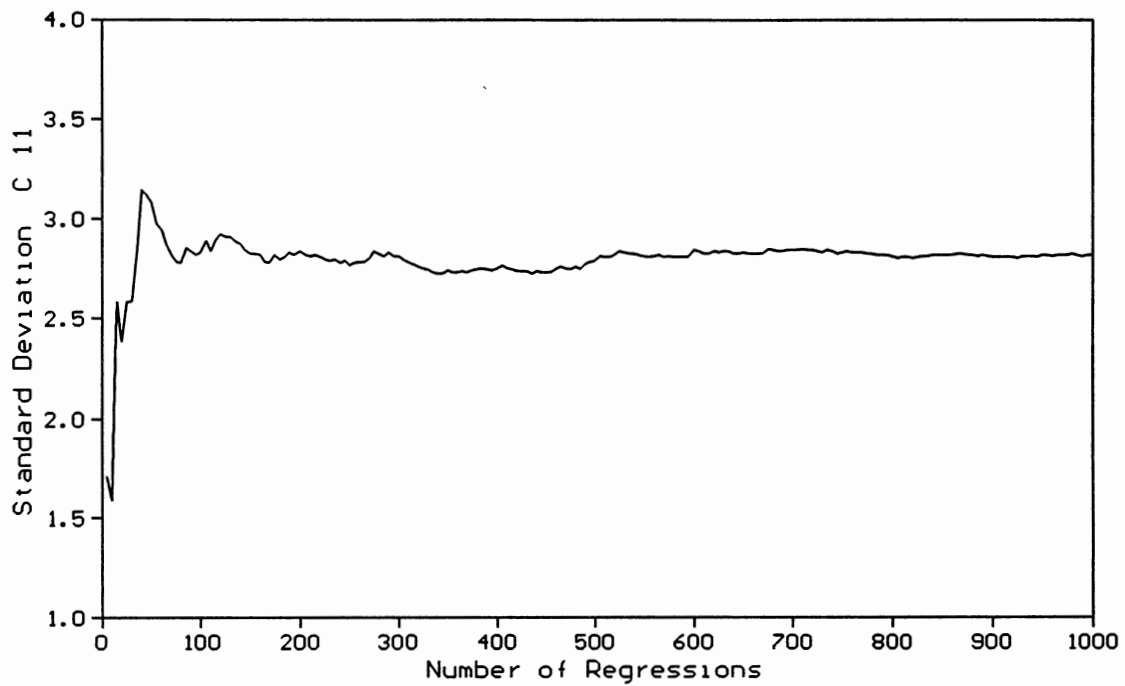


Figure E.22. Stochastic Convergence of Std C₁₁

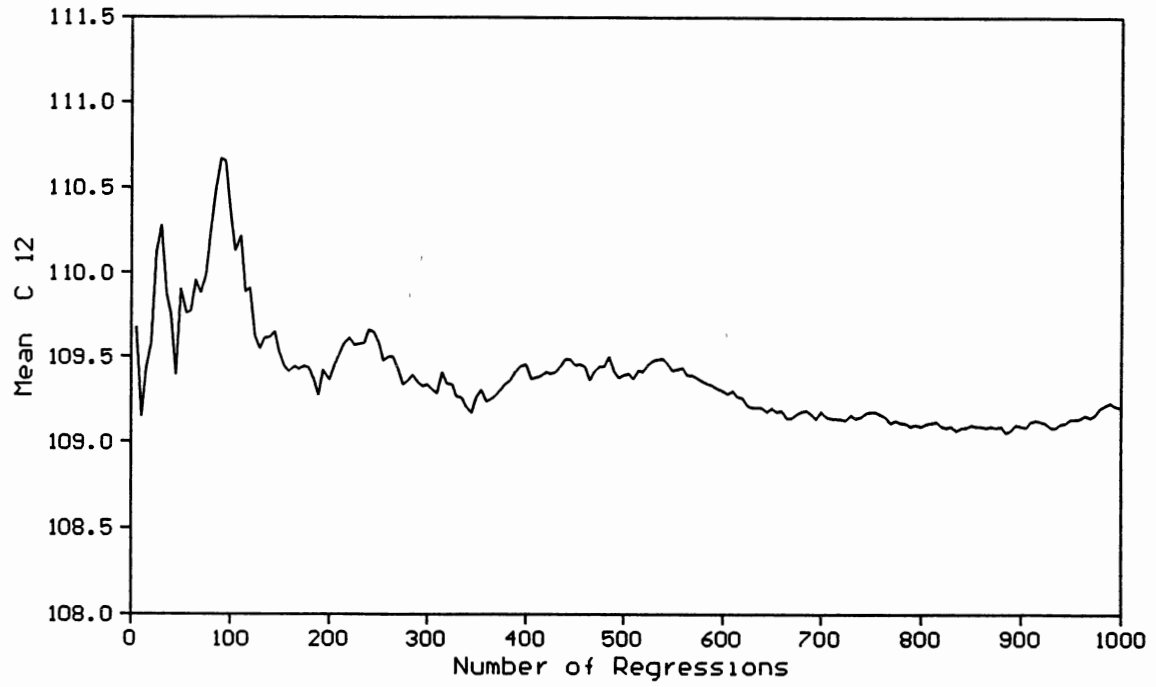


Figure E.23. Stochastic Convergence of Mean C₁₂



Figure E.24. Stochastic Convergence of Std C₁₂

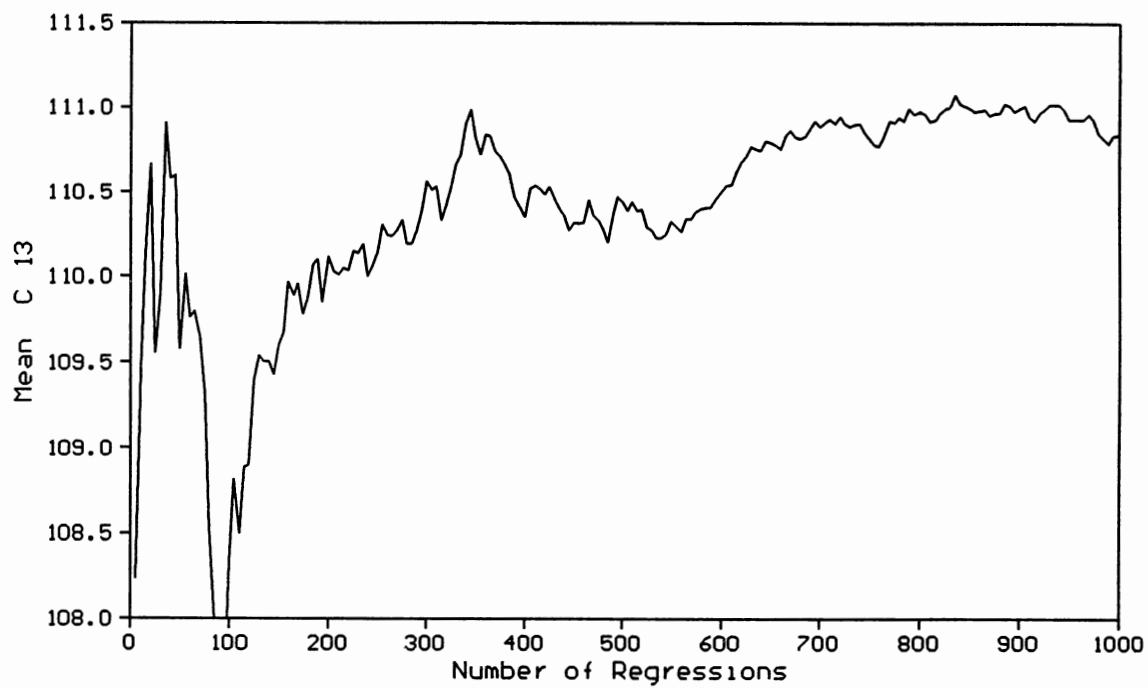


Figure E.25. Stochastic Convergence of Mean C₁₃

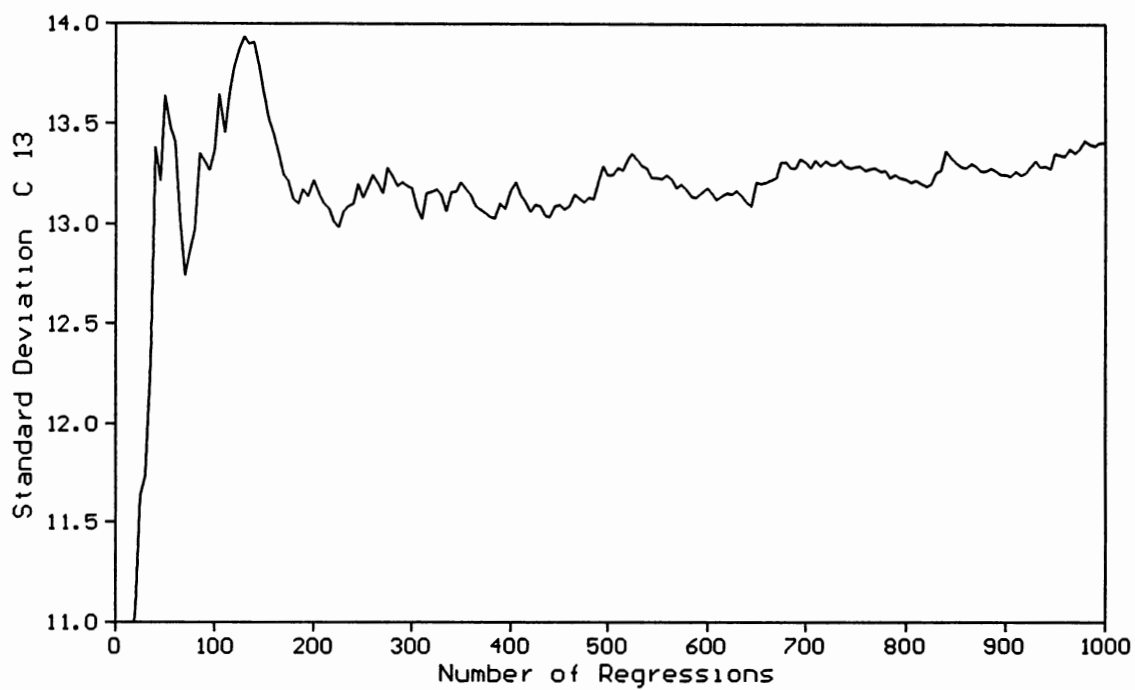


Figure E.26. Stochastic Convergence of Std C₁₃

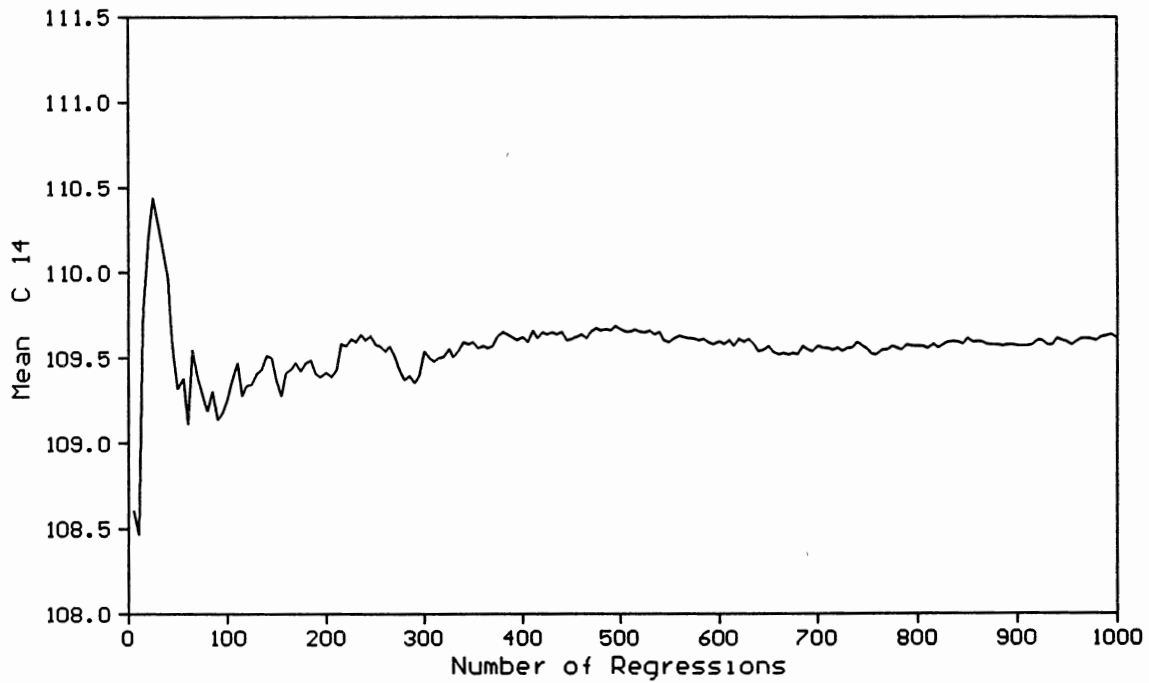


Figure E.27. Stochastic Convergence of Mean C_{14}

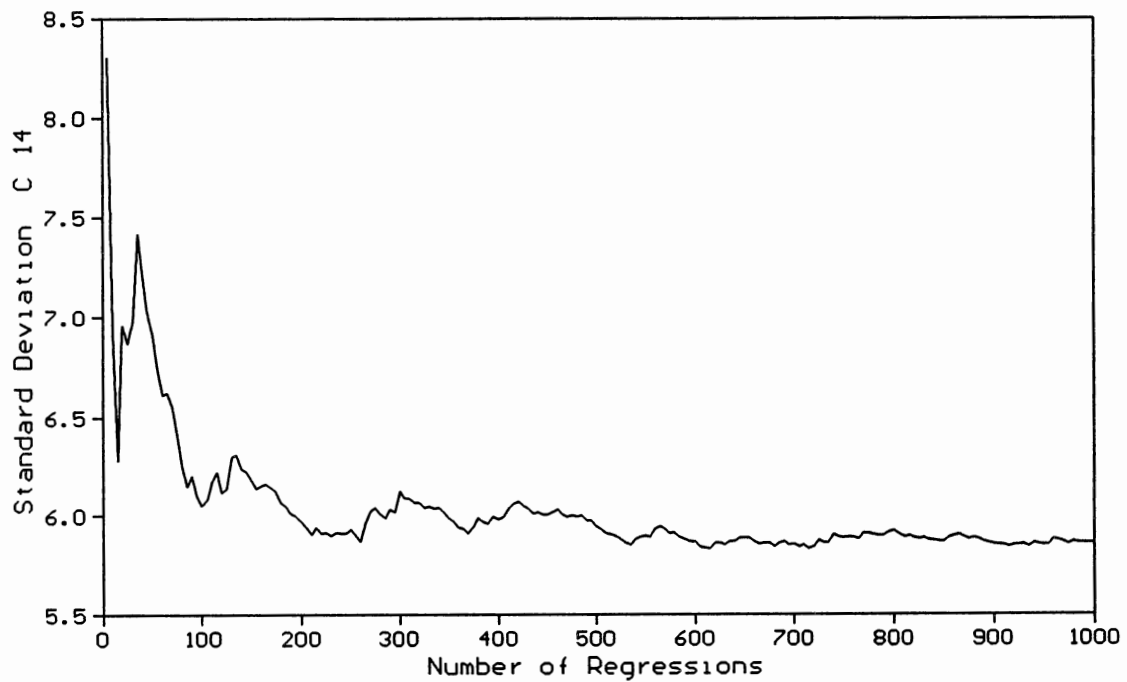


Figure E.28. Stochastic Convergence of Std C_{14}



Figure E.29. Stochastic Convergence of Mean C₁₅

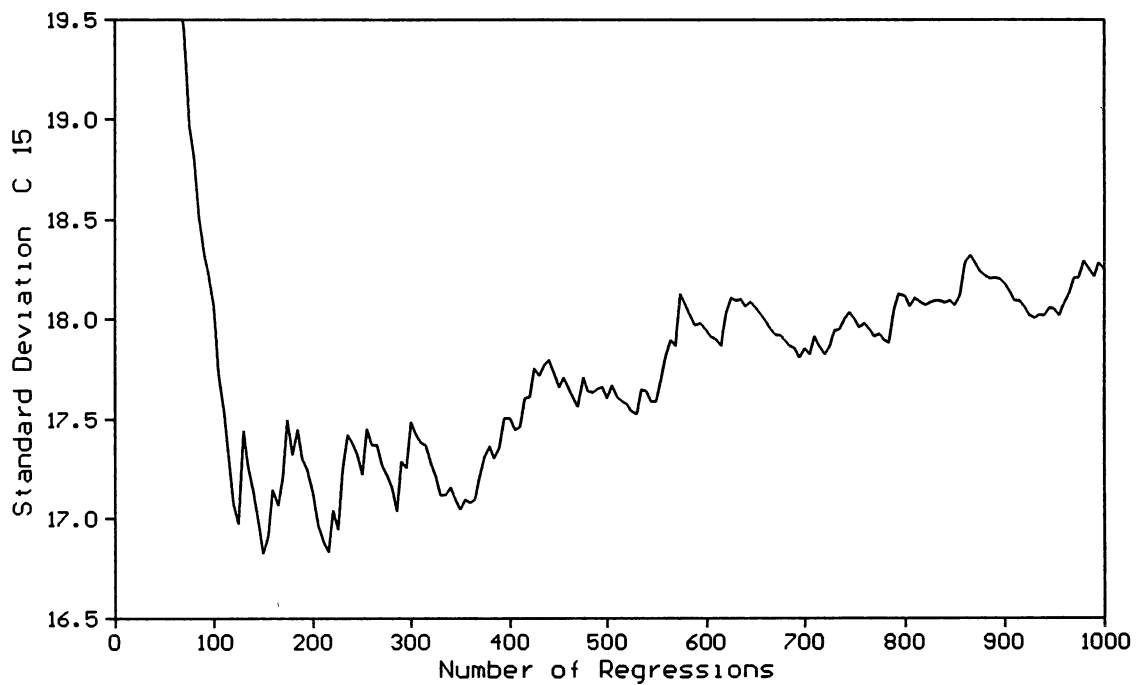


Figure E.30. Stochastic Convergence of Std C₁₅

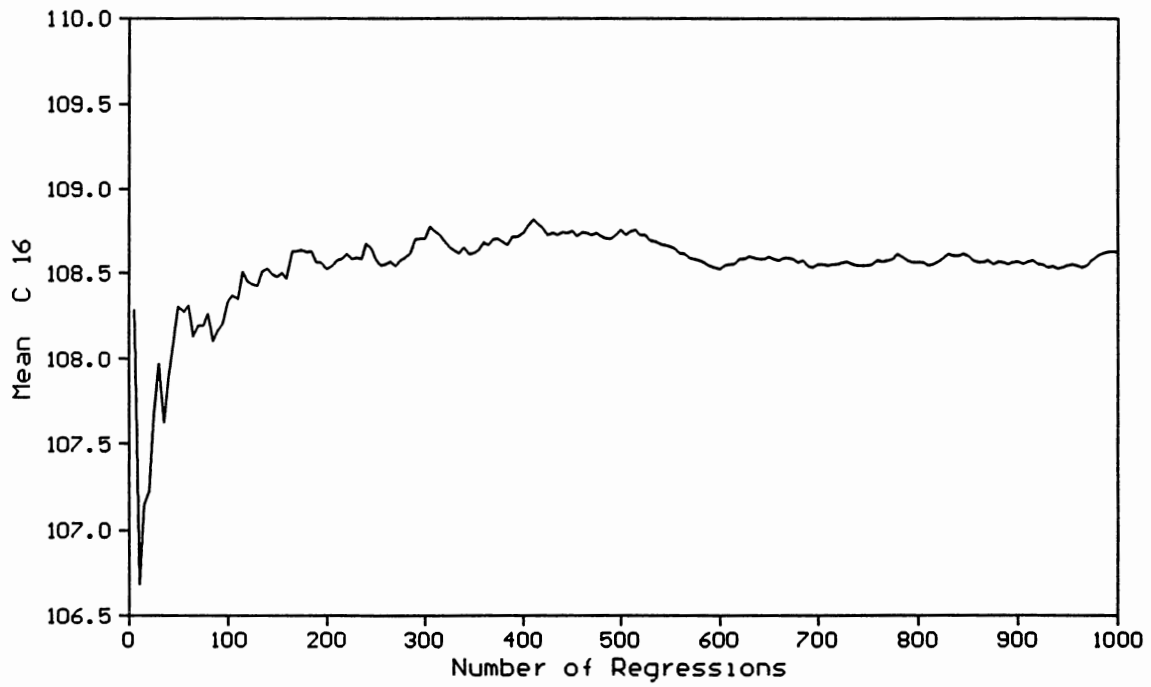


Figure E.31. Stochastic Convergence of Mean C_{16}

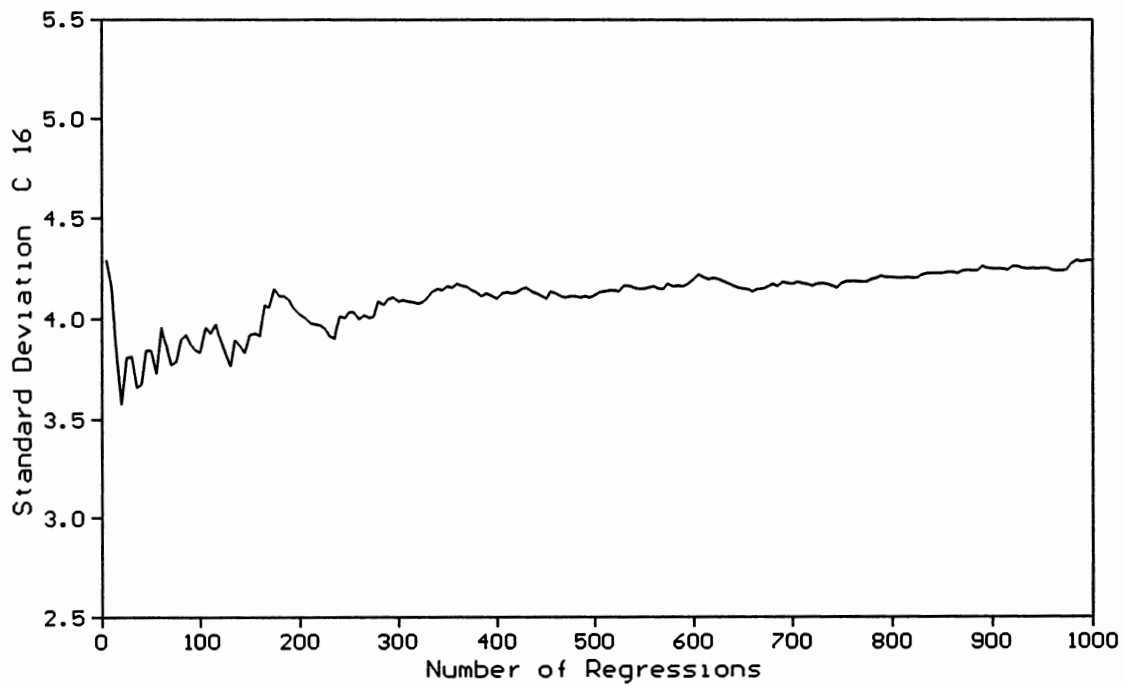


Figure E.32. Stochastic Convergence of Std C_{16}

APPENDIX F

PROBABILITY DISTRIBUTION OF ESTIMATED
ROUGHNESS PARAMETERS

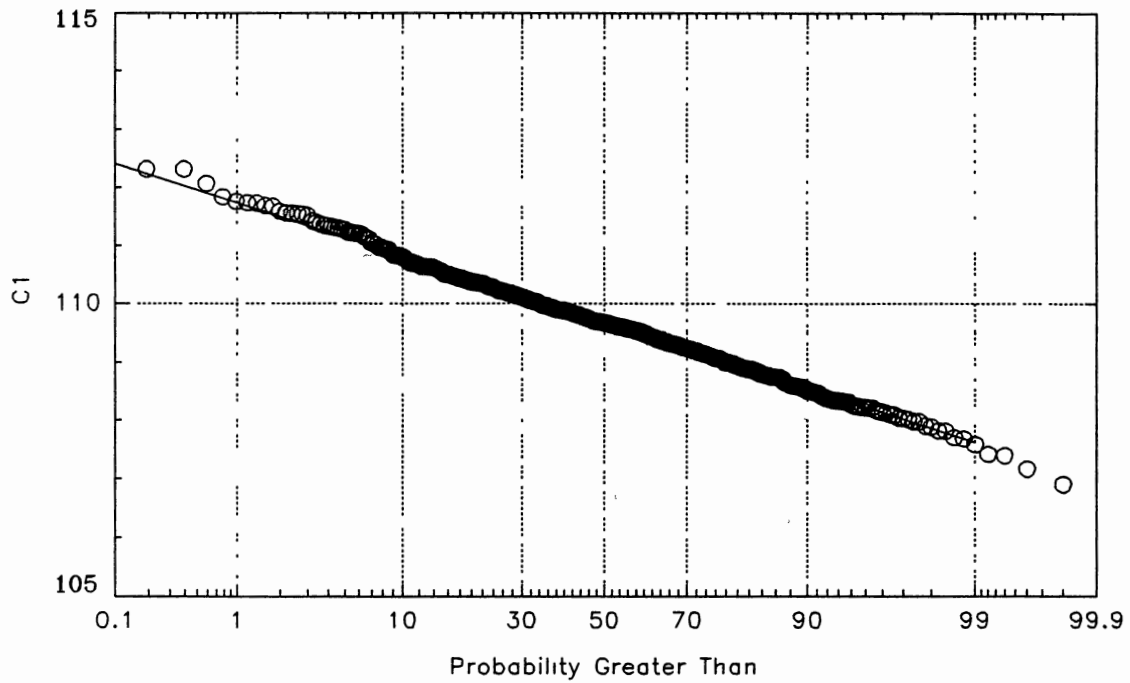


Figure F.1. Probability Distribution of C_1

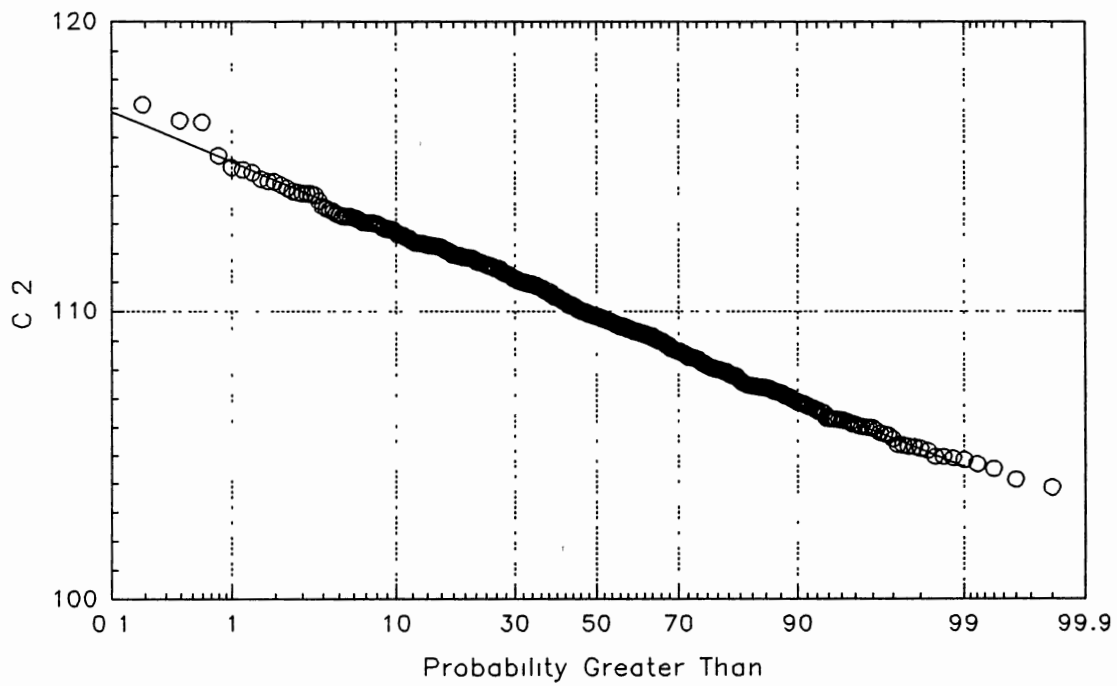


Figure F.2. Probability Distribution of C_2

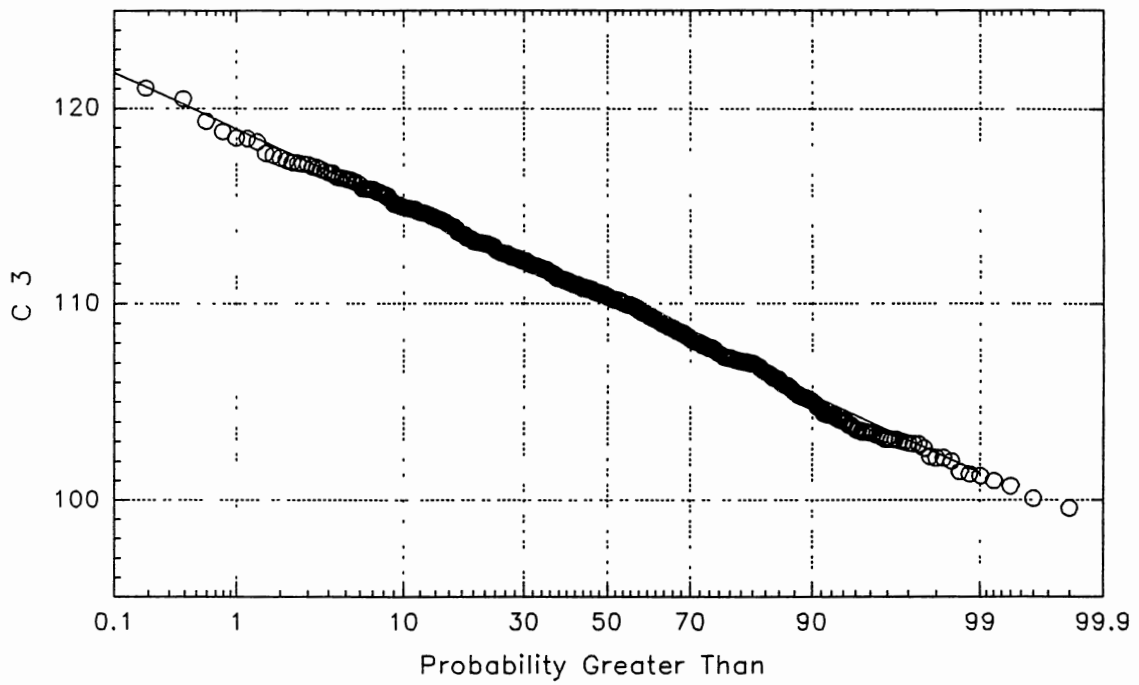


Figure F.3. Probability Distribution of C_3

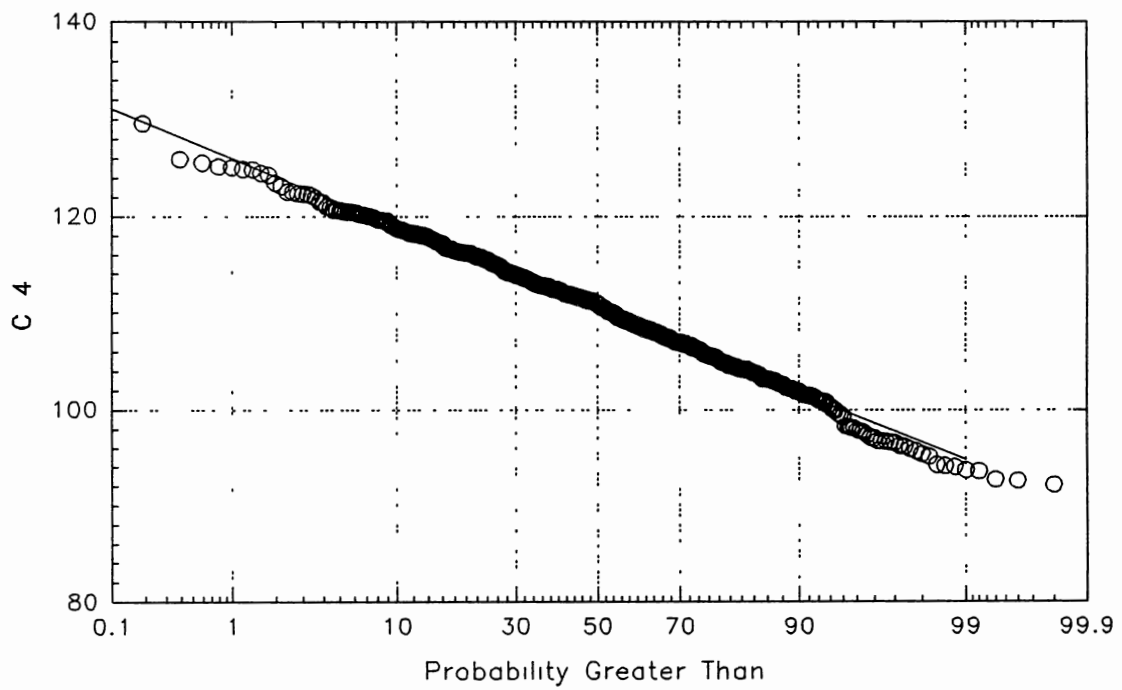


Figure F.4. Probability Distribution of C_4

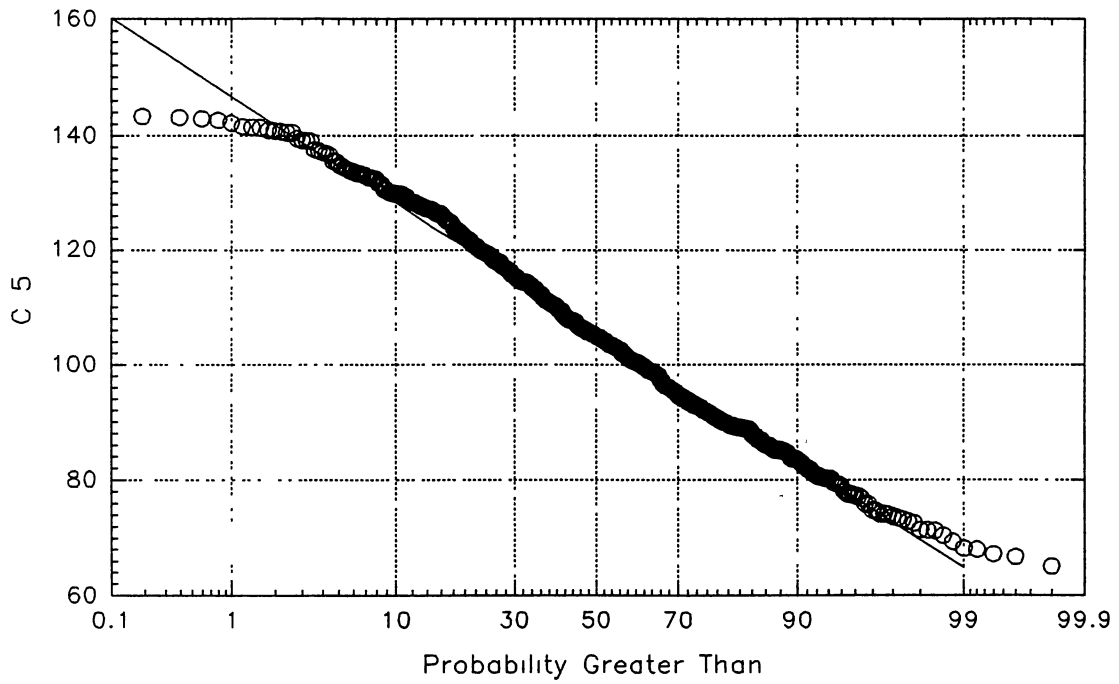


Figure F.5. Probability Distribution of C_5

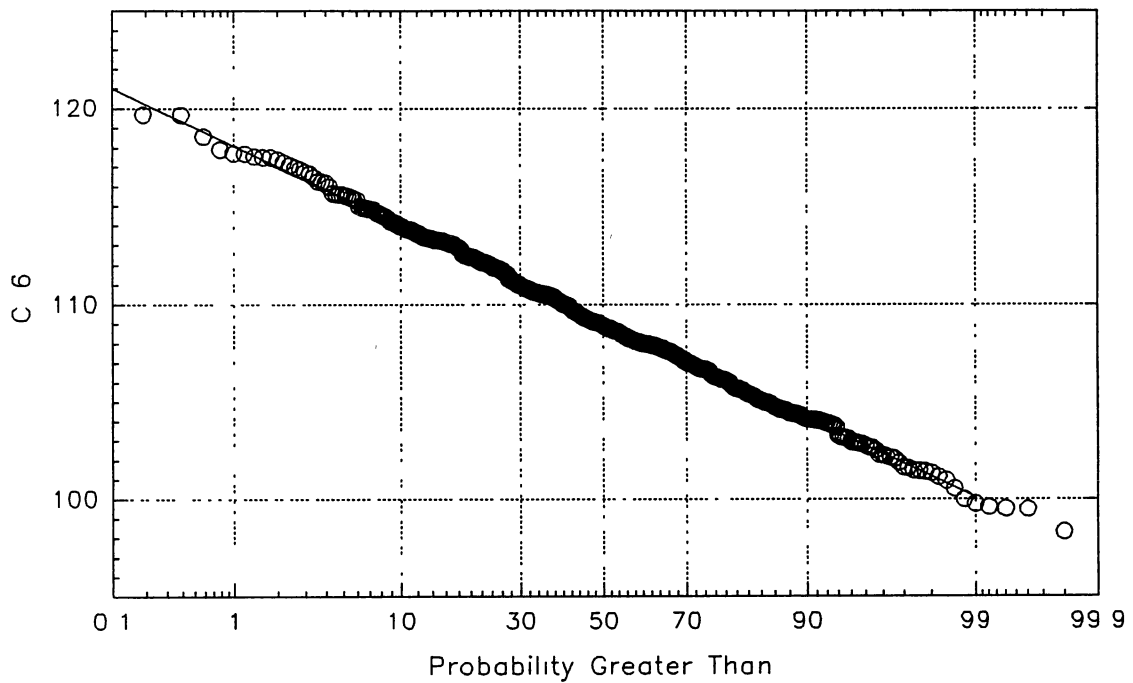


Figure F.6. Probability Distribution of C_6

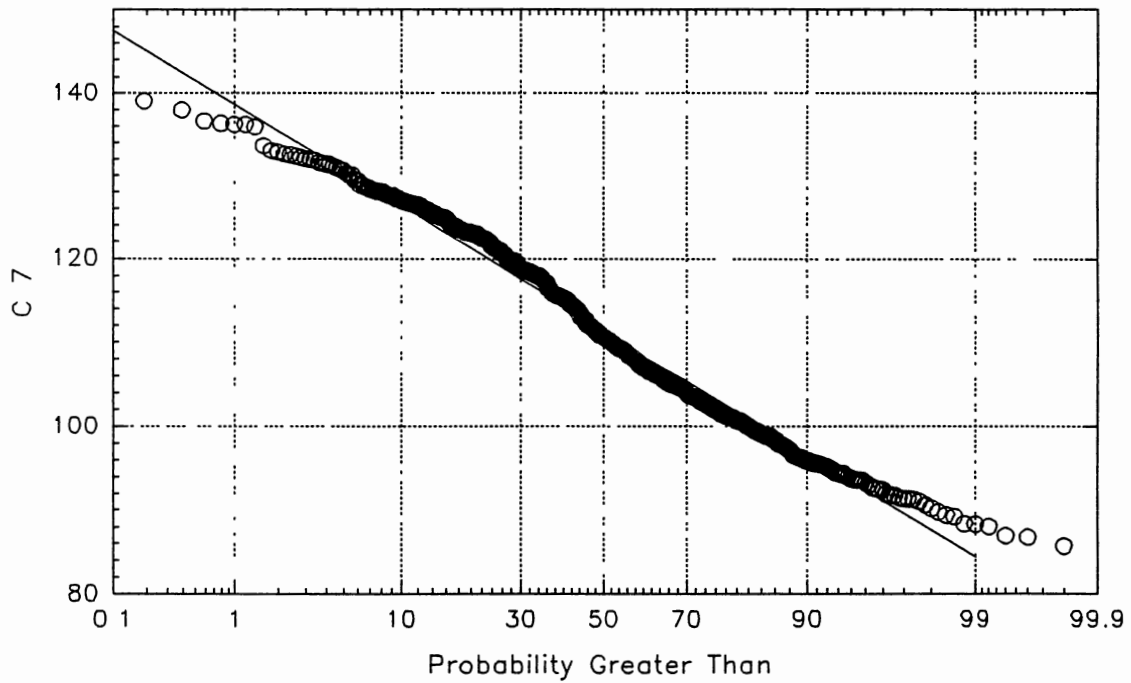


Figure F.7. Probability Distribution of C_7

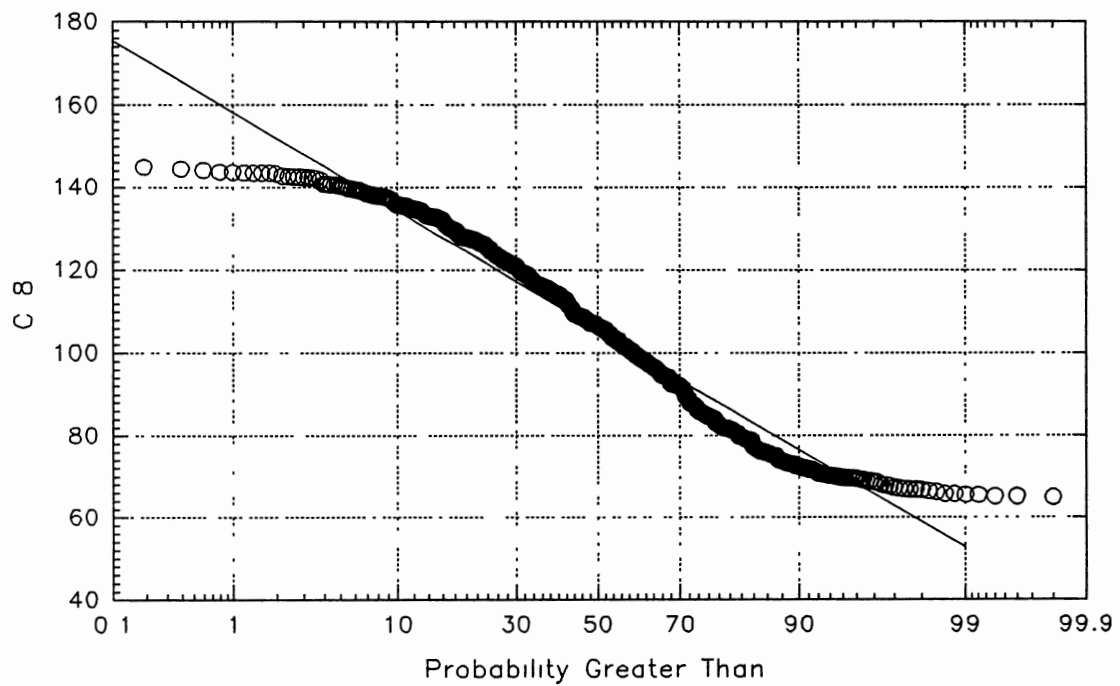


Figure F.8. Probability Distribution of C_8

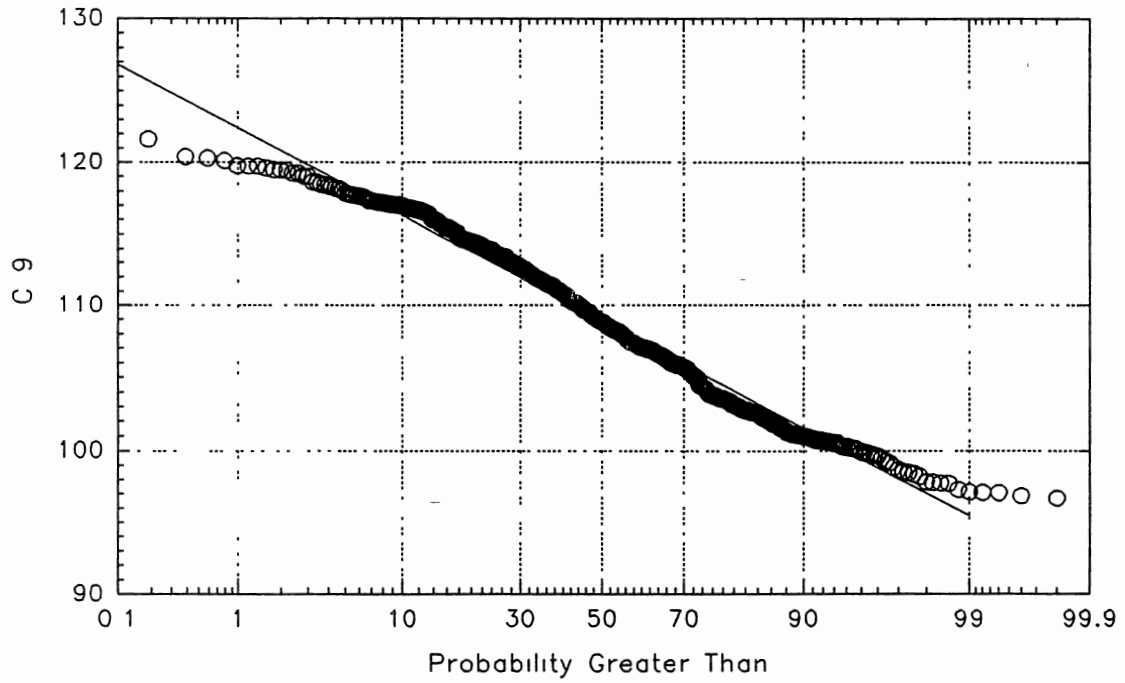


Figure F.9. Probability Distribution of C_9

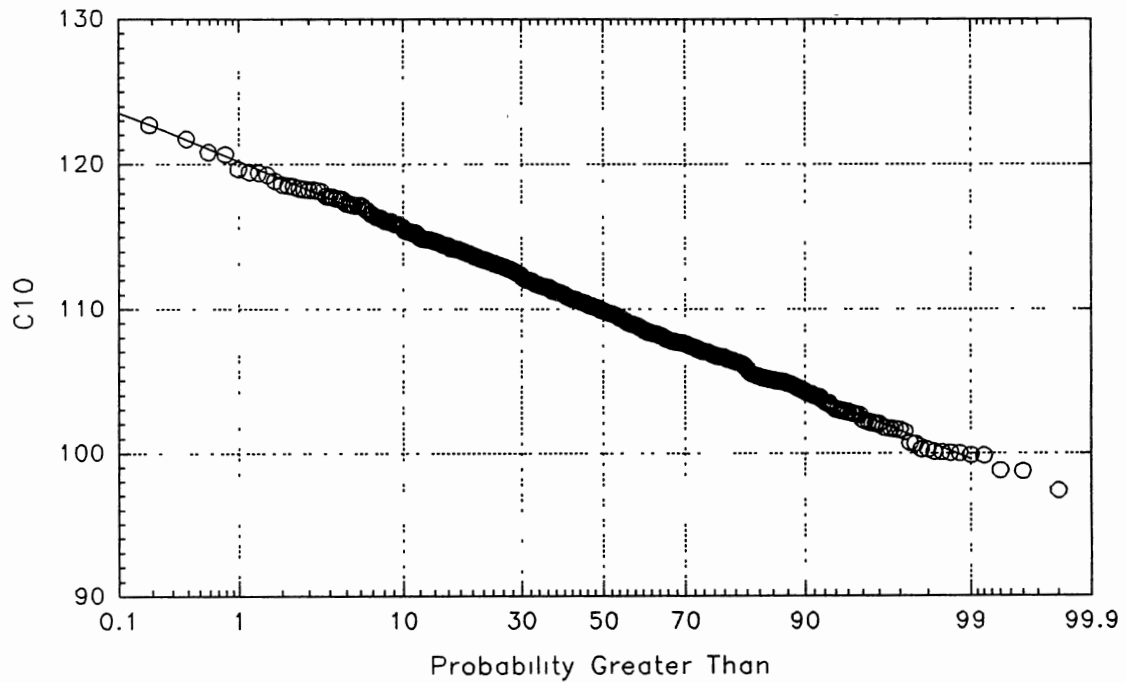


Figure F.10. Probability Distribution of C_{10}

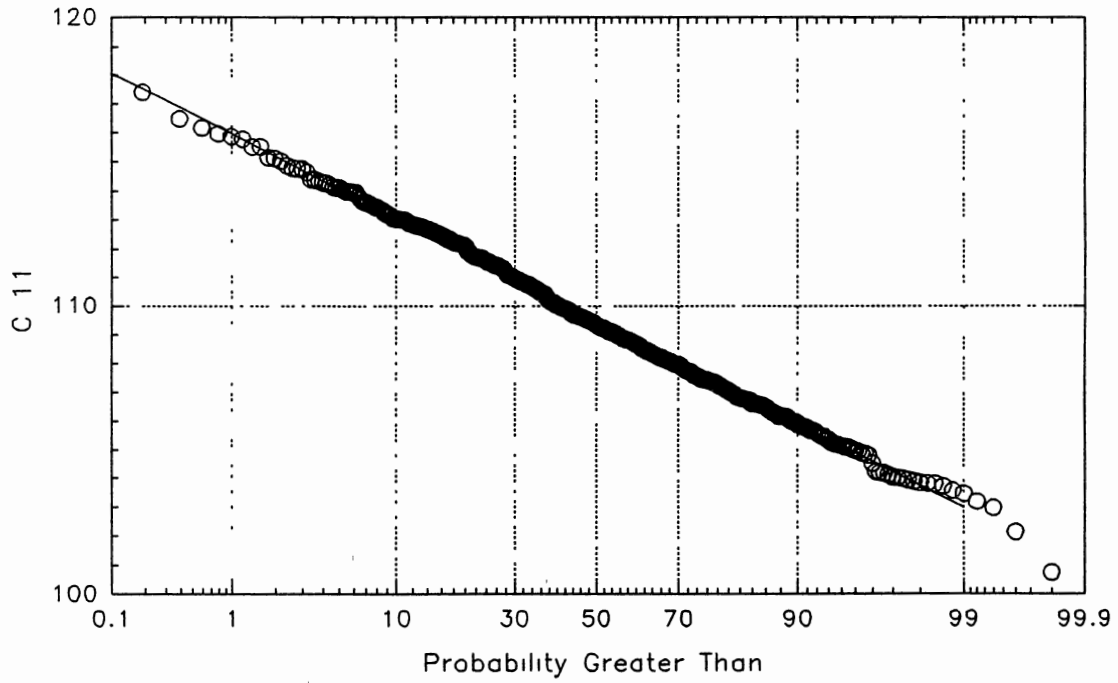


Figure F.11. Probability Distribution of C_{11}

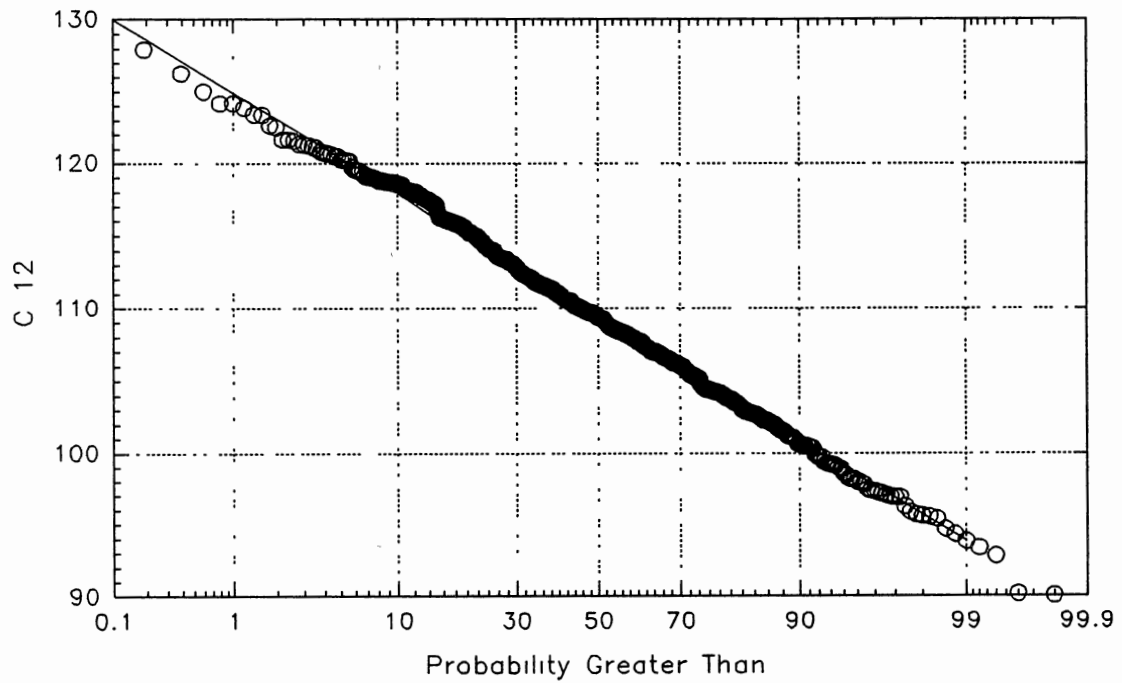


Figure F.12. Probability Distribution of C_{12}

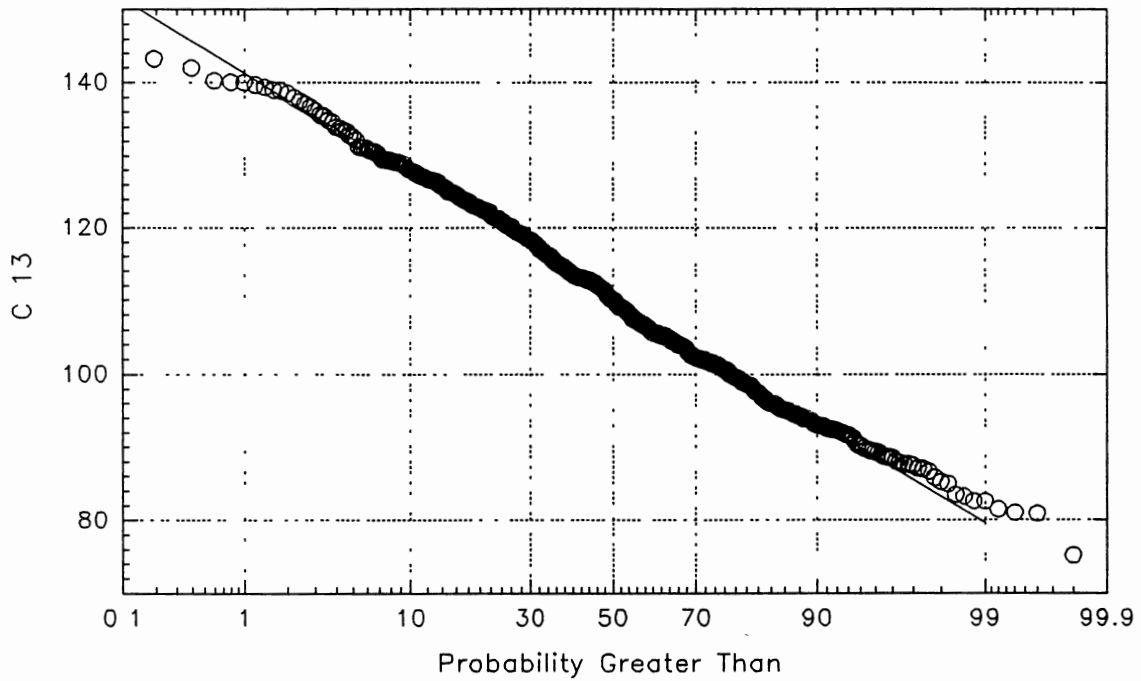


Figure F.13. Probability Distribution of C_{13}

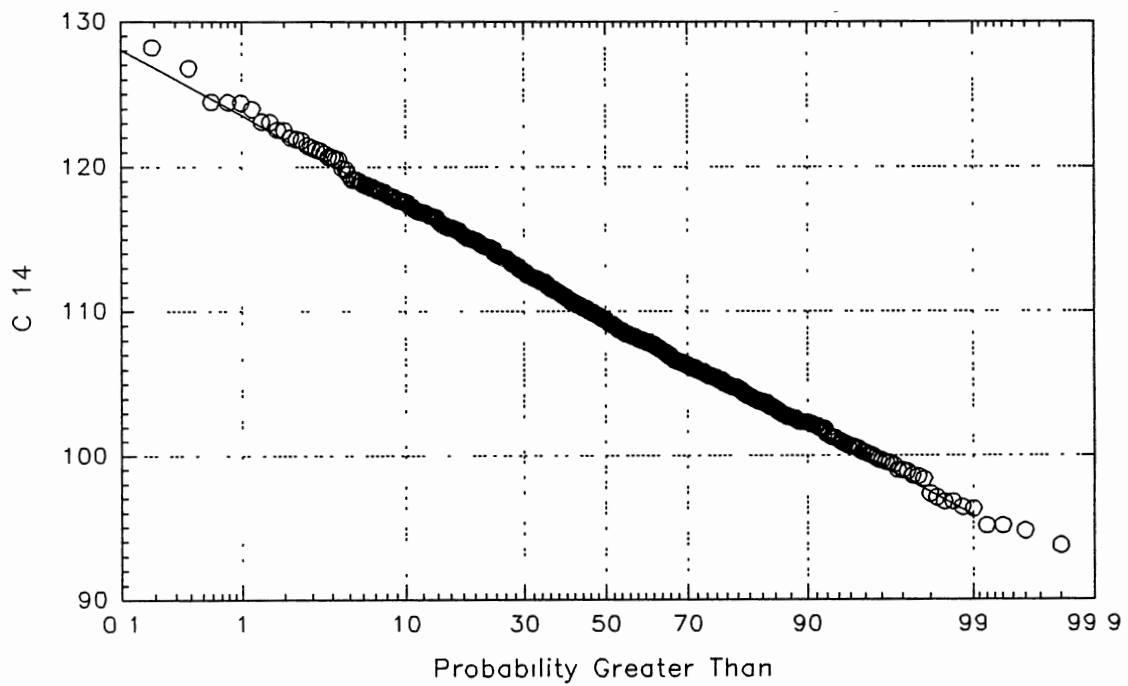


Figure F.14. Probability Distribution of C_{14}

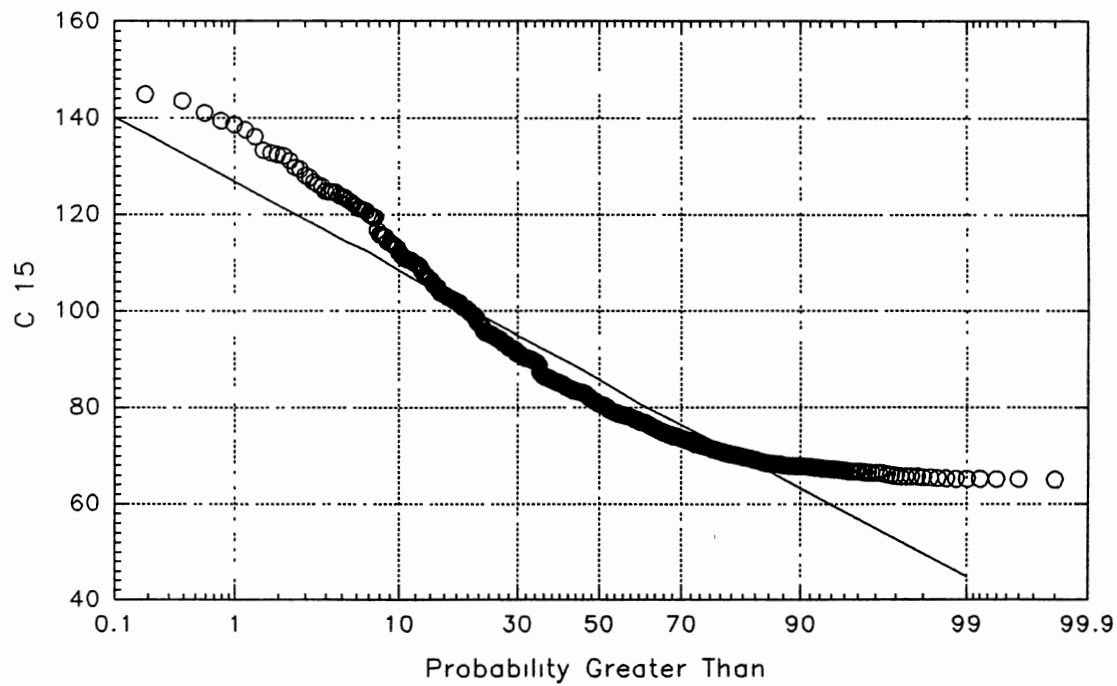


Figure F.15. Probability Distribution of C_{15}

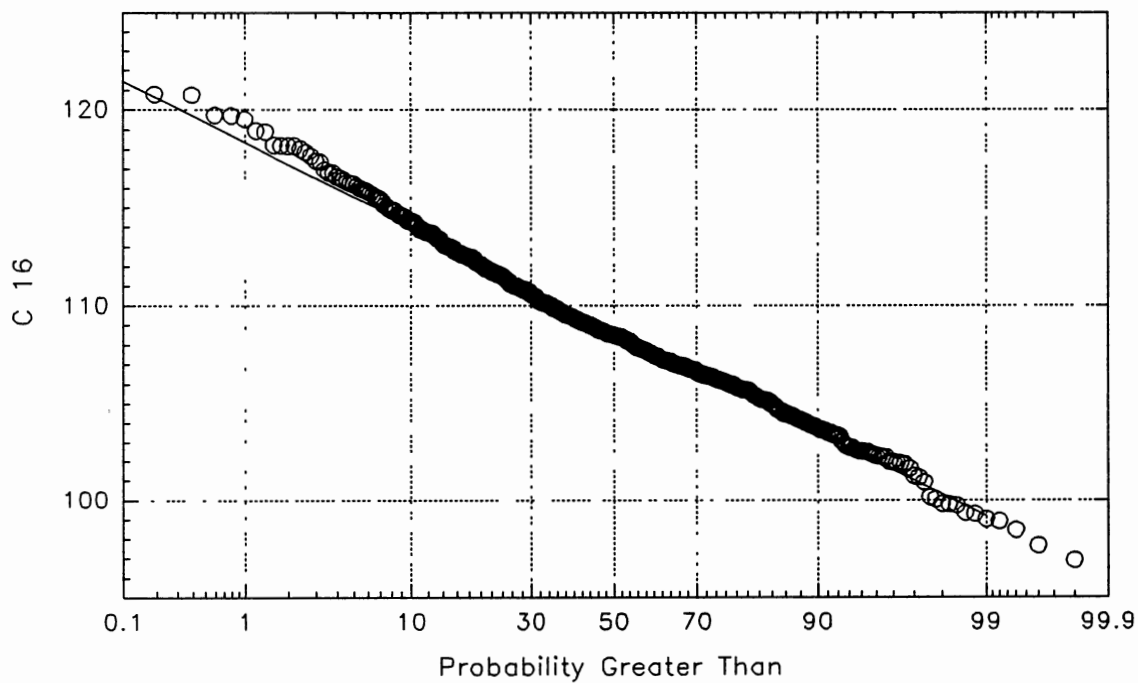


Figure F.16. Probability Distribution of C_{16}

APPENDIX G
PROBABILITY DISTRIBUTION OF
COMPUTED NODAL PRESSURES

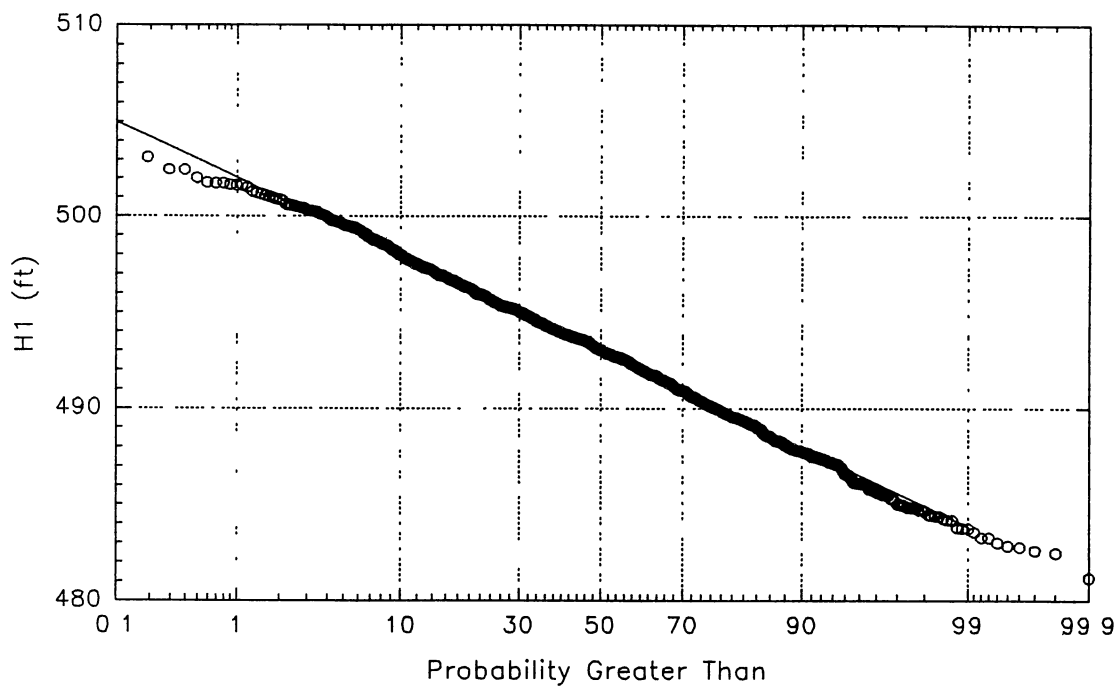


Figure G.1. Probability Distribution of H_1

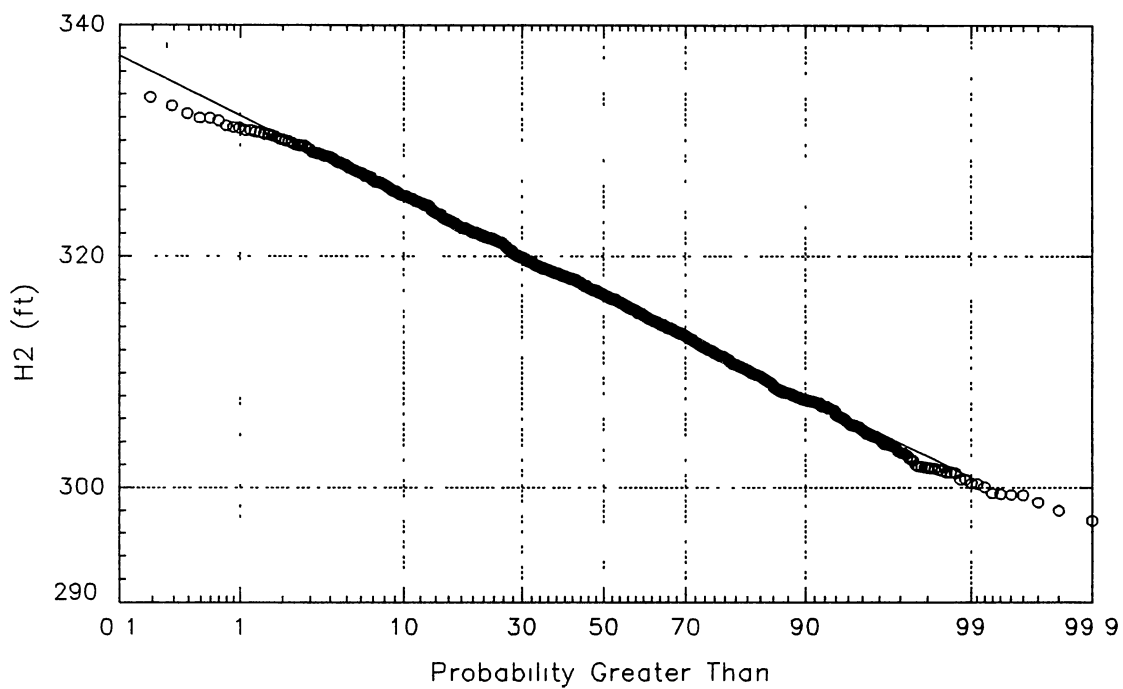


Figure G.2. Probability Distribution of H_2

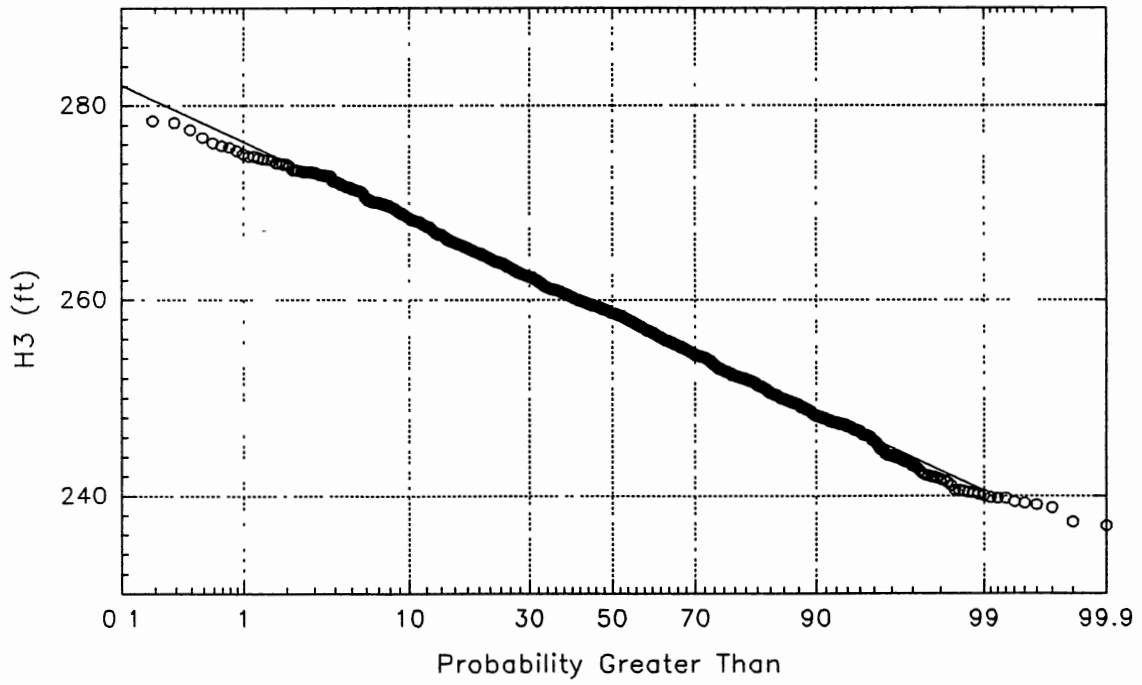


Figure G.3. Probability Distribution of H_3

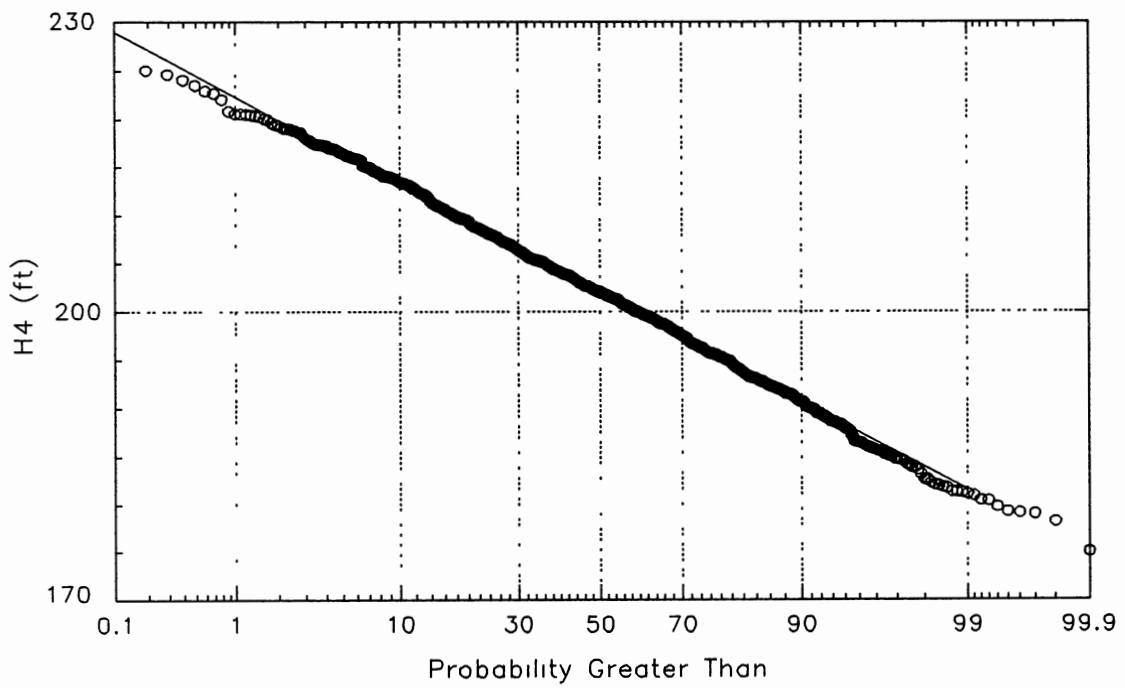


Figure G.4. Probability Distribution of H_4

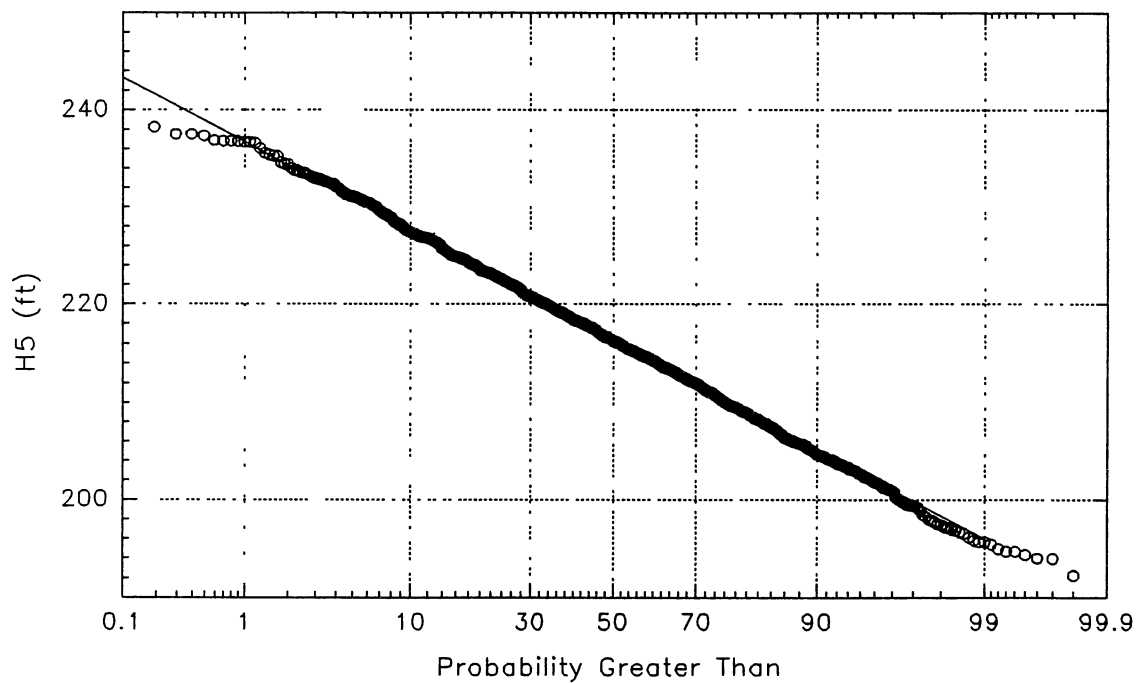


Figure G.5. Probability Distribution of H_5

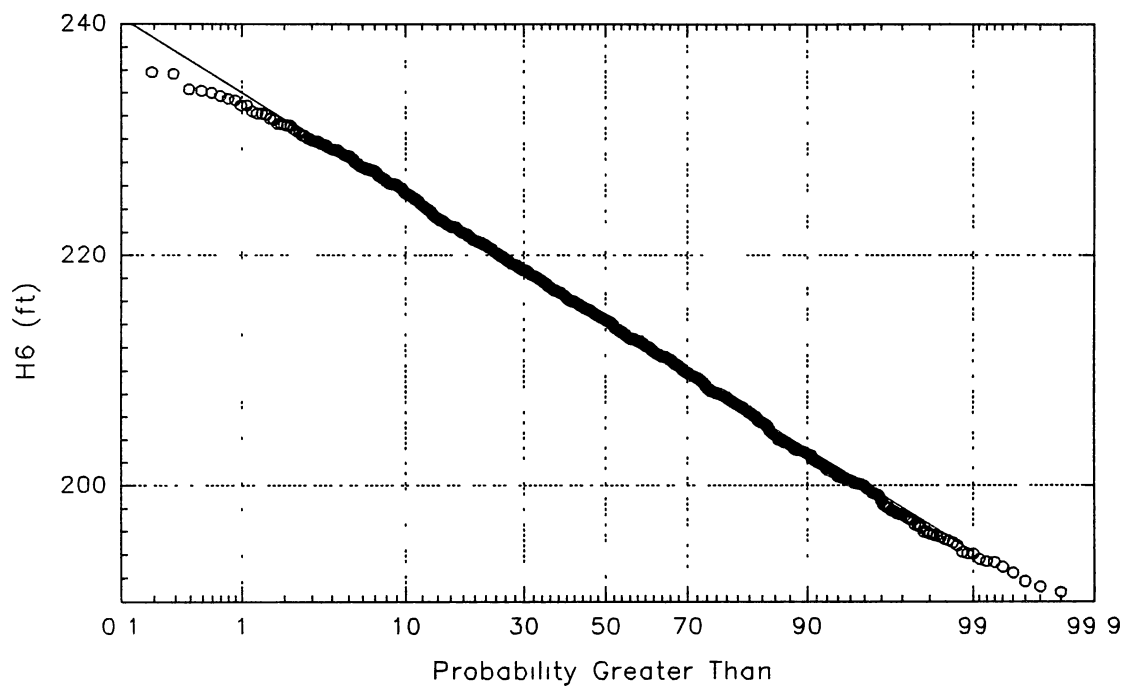


Figure G.6. Probability Distribution of H_6

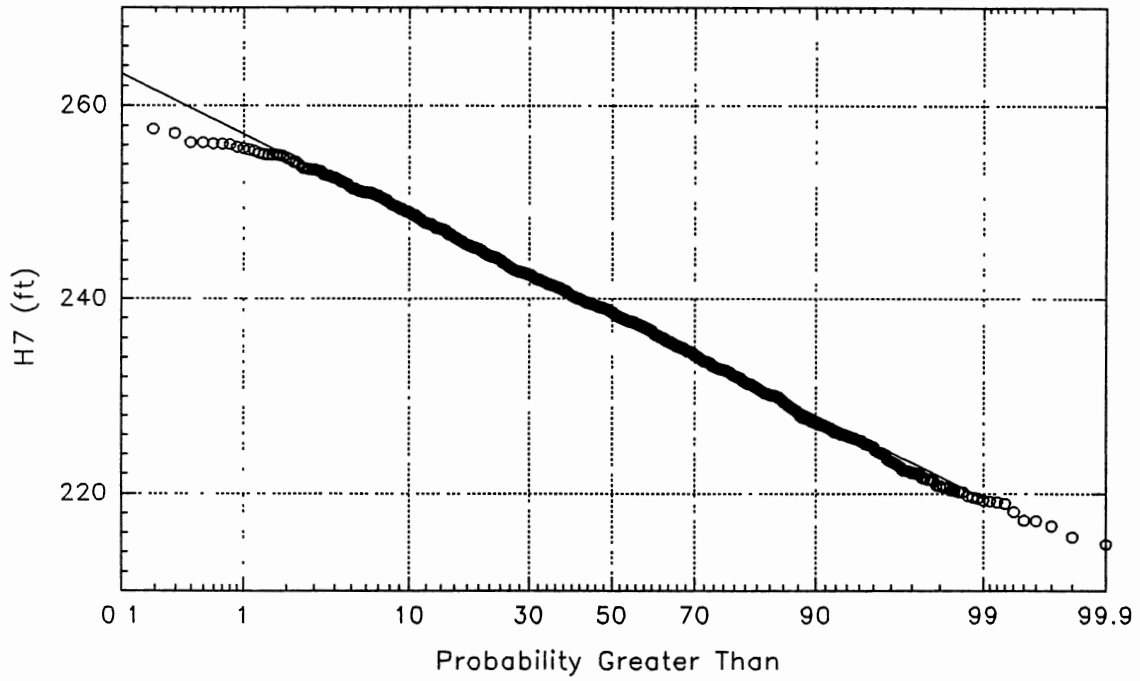


Figure G.7. Probability Distribution of H_7

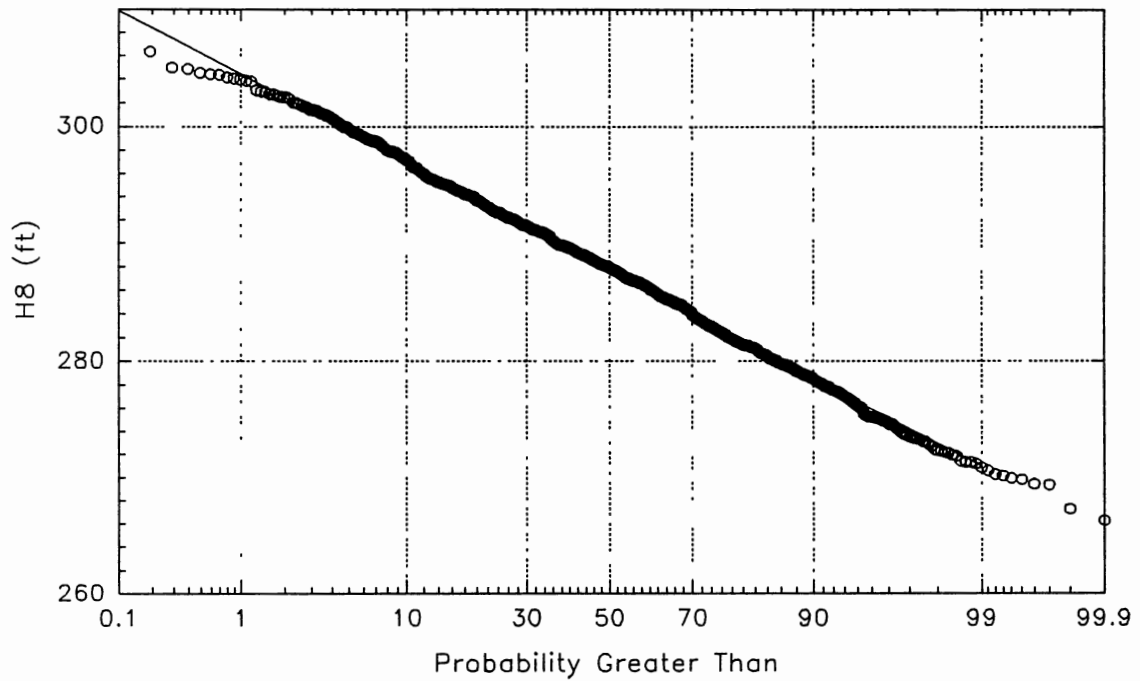


Figure G.8. Probability Distribution of H_8

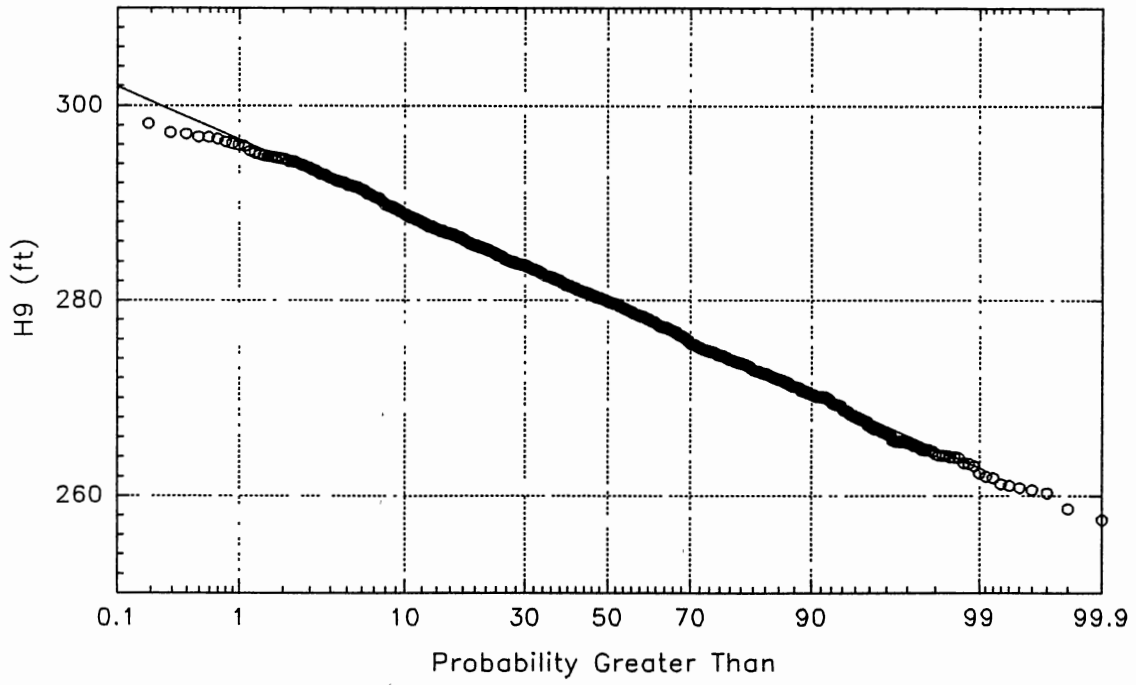


Figure G.9. Probability Distribution of H_9

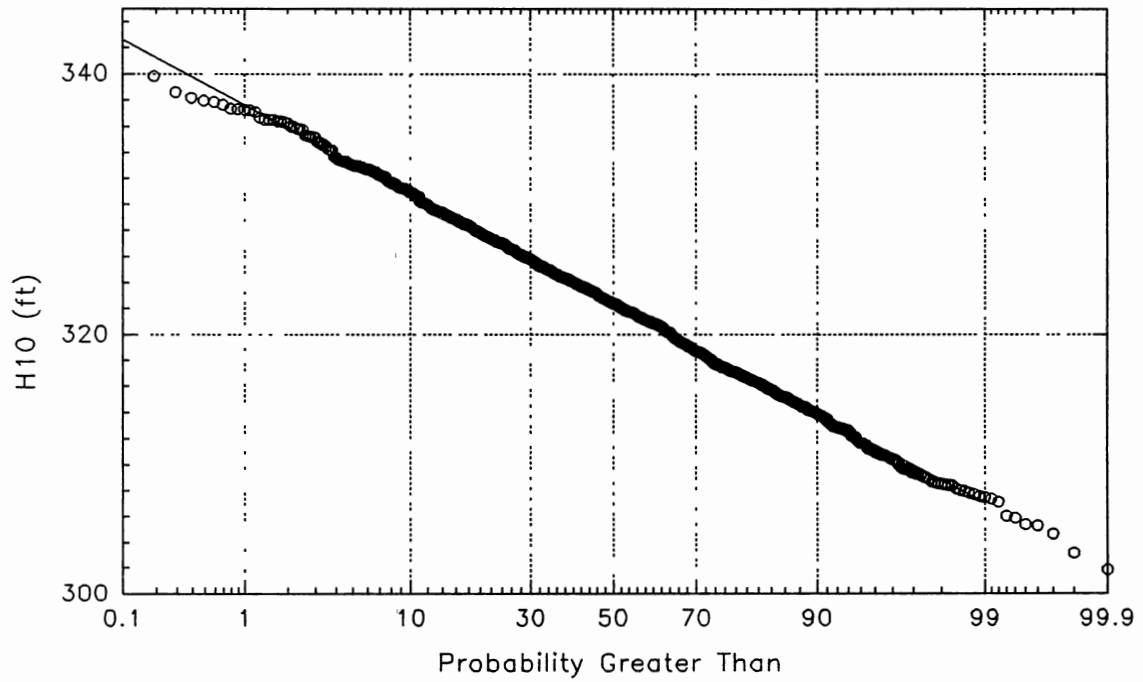


Figure G.10. Probability Distribution of H_{10}

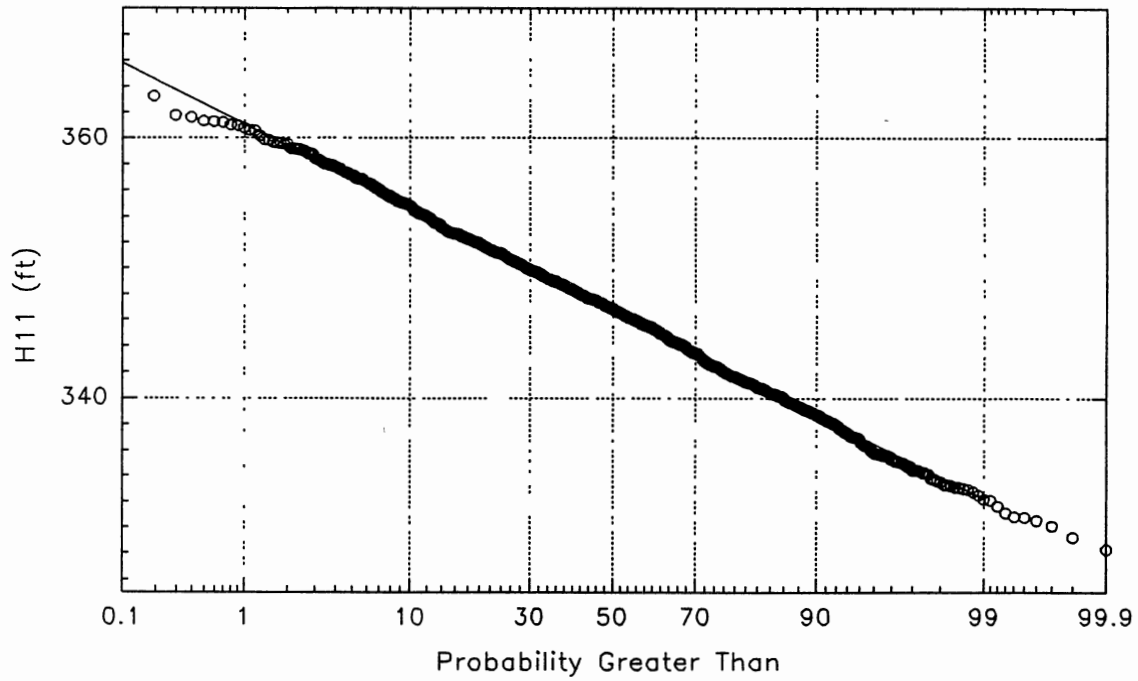


Figure G.11. Probability Distribution of H_{11}

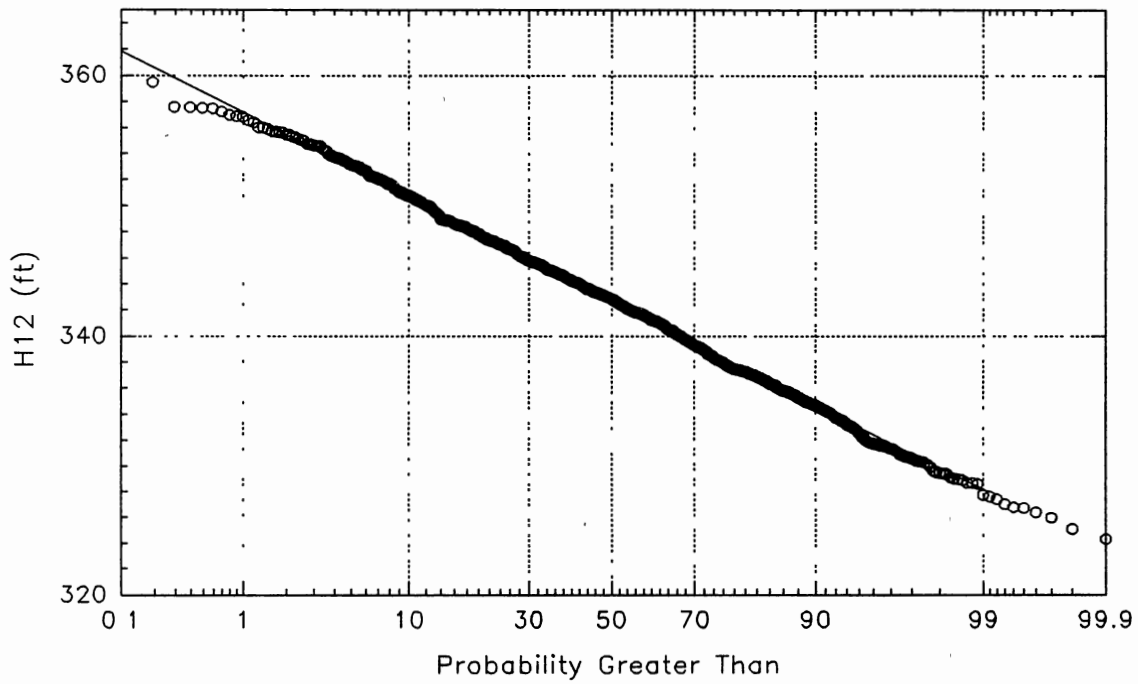


Figure G.12. Probability Distribution of H_{12}

2
VITA

José Vicente Granato de Araújo

Candidate for the Degree of

Doctor of Philosophy

Thesis: A STATISTICALLY BASED PROCEDURE FOR CALIBRATION OF
WATER DISTRIBUTION SYSTEMS

Major Field: Civil Engineering

Biographical:

Personal Data: Born in São Paulo, S.P., Brazil,
September 25, 1958, the son of Djalma Barros de
Araújo and Isa Itália Granato de Araújo.

Education: received Bachelor of Science Degree in
Civil Engineering from Universidade Federal de
Goiás, Goiânia, Brazil, in August 1981; received a
specialization degree in Hydrology from the
University of Padova, Monselice, Italy, in August
1984; received the Master of Science Degree in
Civil Engineering from Oklahoma State University
in May 1989; completed the requirements for the
Doctor of Philosophy degree at OSU in May, 1992.

Professional Experience: Teaching Assistant, Oklahoma
State University, January 1990 to August 1991;
Graduate Research Assistant, OSU, August 1987 to
August 1988 and January 1989 to August 1990;
Assistant Professor of the Department of Civil
Engineering, Universidade Católica de Goiás,
Goiânia, Goiás, Brazil, March to December 1986;
Teaching Assistant, Department of Computer
Science, UFG, Goiânia, Goiás, Brazil, August 1985
to December 1986; Technical Assistant, XXI
International Course in Hydrology, University of
Padova, Monselice, Italy, January to August 1985;
Design Engineer, Water Supply Company of Goiás,
from 1981; Computer Analyst, CPD, UFG, from 1983
to 1989; Design Engineer, Prisma Projetos e
Consultoria Ltda, Goiânia, Goiás, Brazil, 1981 to
1983.

Membership in Professional Societies: American Society
of Civil Engineers, American Water Resources
Association, Water Environmental Federation,
Associação Brasileira de Engenharia Sanitária,
Clube de Engenharia de Goiás.

Membership in Honorary Societies: Chi Epsilon, Tau
Beta Pi, Phi Kappa Phi.