UNIVERSITY OF CINCINNATI

July 21 , 20 03

I, Gayatri Nadimpalli ,
hereby submit this as part of the requirements for the degree of:

Master of Science

in:

Environmental Engineering

It is entitled:

Estimating Leaks in Water Distribution Systems by Sequential

Statistical Analysis of Continuous Flow Readings

Approved by:

Dr. Steven G. Buchberger

Dr. Dionysios D. Dionysiou

Dr. Robert M. Clark

Estimating Leaks in Water Distribution Systems by Sequential Statistical Analysis of Continuous Flow Readings

A thesis submitted to the

Division of Research and Advanced Studies
of the University of Cincinnati

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in the Department of Civil and Environmental Engineering
of the College of Engineering

2003

by

Gayatri Nadimpalli

B.E., Osmania University, Hyderabad, India, 2001

Committee Chairman: Dr. Steven G. Buchberger

ABSTRACT

ESTIMATING LEAKS IN WATER DISTRIBUTION SYSTEMS BY SEQUENTIAL STATISTICAL ANALYSIS OF CONTINUOUS FLOW READINGS

By

Gayatri Nadimpalli

Committee Chairman: Dr. Steven G. Buchberger

Leakage in water distribution pipes is a major problem faced by the water industry. Water utilities often employ traditional audit methods to estimate water lost as leakage. Many hydraulic models have also been developed in the recent past to estimate leakage rates and locate leaks. However, water audits give an approximate estimate of the leakage rates and the mathematical models can be applied under certain hydraulic conditions only. In this thesis, a new method is presented for detecting the magnitude of leaks in residential service zones of a drinking water distribution system. It is assumed that continuous measurements of flow rates through the main supply line into a residential service zone are available during periods of low use. The procedure involves computing the sample mean and variance from the set of measured flow rates as these flows are truncated progressively from below. Trajectories of the sample statistics and their derivatives are plotted versus the level of data truncation. In the presence of leaks, these trajectories diverge from their expected theoretical path when plotted on a standardized graph derived from a mixed truncated normal distribution. The point of departure on the standardized graph indicates where the truncation threshold matches the maximum rate of network leakage. A performance limit for the proposed method is derived to account for spatial constraints reflecting network size and time constraints arising from interval averaging. A simple example is presented first where the leakage is assumed to be of constant magnitude. The leak analysis is then extended to a more complicated case where pressure fluctuations in the distribution system add a statistical noise to the flow measurements. Results show that the method developed in this thesis estimates leakage rates quite accurately even in the presence of statistical noise. The method can be used to estimate leakage rates in branching mainlines or residential District Metering Areas (DMAs) where flows can be measured continuously for a sufficiently long period of time.

ACKNOLEDGEMENTS

I would like to thank my advisor, Dr. Steven Buchberger for his excellent guidance, encouragement and support. Working as a research assistant under Dr. Buchberger has taught me good organization presentation skills, which will be very useful throughout my career. Thank you Dr. Buchberger!

I would also like to thank Dr. Robert M. Clark and Dr. Dionysiou for their encouragement and interest in my thesis work.

I am grateful to my senior and co-researcher Zhiwei Li, for his support with the Poisson Rectangular Pulse Simulator and advice his during the two years of academic career at the Department of Civil and Environmental Engineering, UC. Frequent discussions with Zhiwei Li have helped me in improving my thinking ability.

I am also grateful to my parents for the constant support and encouragement they have given me right from my childhood.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbol	Meaning	Units
DMA	demand monitoring area	
PRP	Poisson rectangular pulse	
b	coefficient in leak simulation algorithm	
K	number of arrivals	/min/home
m	rank of measured flows in descending order	
M	number of flow data points in the upper tail region	
N	number of homes in monitoring area	
p	p^{th} quantile of standard normal distribution	
P(0)	stagnation probability	
Q_L	leakage rate	L/min
Q	flow rate without leakage	L/min
Q*	rescaled flow rate	
y'	slope of sample moment	
Z	standard normal variate	
q^l	leak quantity	m^3/sec
A^l	combined coefficient for discharge and leak area	m^2
p	pressure	N/m^2
g	gravitational acceleration	m/sec^2
E[Q]	mean of flow	L/min
$E[Q^2]$	square of mean of flow	$(L/min)^2$
$E'[Q^*]$	slope of rescaled mean	
S[Q]	standard deviation of flow	L/min
$S'[Q^*]$	slope of the rescaled standard deviation	
V[Q]	variance of flow	$(L/min)^2$
Δt	time step	sec
r(1)	lag-1 autocorrelation coefficient for the leakage rate	
σ	standard deviation of parent normal distribution	L/min
μ	mean of parent normal distribution	L/min
ε	standardized truncation level	
ε_L	point of departure between observed and standardized moments	
λ	mean arrival rate at one home	/min/home
τ	mean pulse duration	min
ϕ(z)	pdf of standard normal distribution	
Φ(z)	cdf of standard normal distribution	
Θ(z)	coefficient of variation	
γ	specific weight of fluid	N/m^3

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND ON PIPE LEAKAGE IN WATER DISRIBUTION NETWORKS

A well-maintained water distribution system is a major asset for any city or community. Continuous monitoring and maintenance of the distribution network is the key step in meeting pressure and flow requirements, and water quality standards. A recent drinking water infrastructure needs survey (USEPA 2001) estimates an investment of 151 billion dollars over a period of 20 years to provide safe drinking water for US customers. Reducing water losses in pipe networks can minimize the maintenance costs and further improve the performance of pipe networks.

Leakage can be defined as unintentional or accidental loss of water from the pipe distribution network (Smith *et al*, 2000). Leaking pipes are a major concern for water utilities around the globe (Table 1.1) as they constitute a major portion of water losses. One of the primary reasons for leakage in pipes is aged and deteriorated networks. The condition of existing old networks can only worsen and further increase water losses. In the US alone, 50% of supplied water is lost as leakage in some of the older networks (Jowitt and Xu, 1990). Leakage rates are also related to length of pipes and number of connections. Improper connections can sometimes result in continuous escape of water from the distribution pipes.

Table 1.1 Network leakage rates around the globe.

Country	Leakage Rate (%)
Netherlands	5
Japan	11
USA	12
France	15
Korea	16
UK	28
India	30

(Source: www3.akwien.at/pdf/uv/university_of_greenwich.pdf).

Other reasons for leakage are unsuitable physical conditions like loading patterns on the ground surface, soil characteristics, temperature, and quality of water carried by pipes (Smith *et al*., 2000). Below freezing temperatures cause contractions in pipes and are considered to be a primary mode of pipe failure in cold countries. Poor water quality may cause corrosion of pipe internal surfaces and contribute to pipe leakage.

Leaks, if undetected, can jeopardize the integrity of pipe networks. Leaking pipes cause an imbalance in the system demands as the supplied water has to be rationed across the network to meet additional demands due to leakage. Further, leaks in distribution pipes lead to serious operational problems, decrease network pressures, inflate energy expenditures, and compromise water quality.

Leak detection techniques that are in use in the water industry involve two major steps.

1. Estimation of leakage rates
2. Location of leaks

1.2 WATER AUDITS

Unaccounted-for water accounts for authorized unmetered use and under-registering metered use in addition to water lost as leakage (Smith *et al.*, 2000). An overall survey of leaks is done as a part of water audit since, leakage is considered as a major component of unaccounted-for water. A water audit involves comprehensive accounting of the total water pumped at a service station and water utilized at the consumer end. The water supplied at a utility is measured while being pumped into the network, and water consumed is obtained from billing records. Water lost as leakage is estimated from simple mass balance principle which can be stated as the difference between the amount of water produced at the utility and amount of water purchased by consumers.

Flow measurements taken in the distribution network give more precise estimates of leakage rates. A portion of the network is isolated using valves and measurements of flows entering the isolated portion of the network are taken for at least a period of 24 hours. Mass conservation principle is applied to that part of the network to estimate the average amount of leakage rate. However, such methods give only an approximate estimate of leakage rates.

1.3 LEAK LOCATION

The second step in leakage control involves location of leaks across the distribution network. Acoustic equipment combined with correlation methods are generally used to locate leaks. Acoustic methods are based on changes in sound of the escaping water (Smith *et al.*, 2000).

Acoustic sound transducers are placed in contact with ground surface to listen to any abnormalities in the underlying pipes. In acoustic equipment accompanied by noise correlators, the acoustic signals from transducers are transmitted to a receiving unit, where the signals are processed automatically. Other leak detection methods employed are infrared thermography, tracer gas methods, and mechanical drilling of soil.

1.4 RESEARCH OBJECTIVES

The primary objective of this research study is to estimate leakage rates from measured pipe flows of a distribution network. The proposed method is an intermediate between simple water audits and sophisticated hydraulic models. The method is based on sequential statistics of continuously measured flows across a distribution network. It is aimed to estimate leakage rates in presence of statistical noise occurring in flow measurements due to fluctuations in the distribution networks. Performance limits are evaluated for application of the method to district metering areas or branching mainlines.

CHAPTER 2

LITERATURE REVIEW

Various mathematical and hydraulic models have been proposed to estimate the magnitude of leakage rates and locate leaks in pipelines and pipe networks. Pudar and Ligget (1992) were able to detect and locate leaks in pipe networks using an inverse problem approach. The method requires pressure and/or flow measurements at various nodal points in the pipe network. The number of measurements are varied to formulate three problems – over-determined, even determined, and underdetermined. Leaks at nodes are considered to be the unknowns and are expressed in terms of pressure by an orifice formula.

$$q = C_o A \sqrt{\frac{2gp}{\gamma}} \tag{2.1}$$

where q = leak at a node

C_o = orifice coefficient

A = equivalent orifice area

p = pressure

g = gravitational acceleration

γ = specific weight of fluid

However, the method can be applied only under steady state conditions, requires accurate pipe friction factors, and huge amount of pressure and flow measurements.

The inverse transient approach adopted by Ligget and Chen (1994) uses transient flow conditions to calibrate and detect leaks in pipe networks. As in the previous work, leaks are defined by an orifice equation. The method is computationally intensive and convergence of solution is not certain.

Liou and Tian (1995) adopted different approaches to detect leaks in petroleum pipelines. According to the Cauchy algorithm, a discrepancy between measured and calculated flows and pressures at the pipe ends declares the presence of a leak. In the presence of noise, the discrepancy between measured and calculated values is overshadowed, and leak detection becomes difficult. Hence, the time-marching algorithm is introduced to avoid noise amplification. Leak location is more accurate in the latter algorithm, as the noise amplification is not carried forward with time. Both algorithms performed well on pipelines simulated with leaks. In the impulse response method (Liou, 1998), a pipeline's response to an impulse in normal conditions and in the presence of leak is discussed. Tests are held to detect leaks by adding noise in the form of pseudo random binary signals. Leaks can be detected and located by this method.

Mukherjee and Narasimhan (1996) used the generalized likelihood ratio test to detect, locate and isolate leaks occurring in pipe networks. Steady state flow conditions are considered and the fluid is assumed to be incompressible. Pressure and flow are the two hydraulic variables measured. An advantage of this method is that a leak can be detected and located even in the pipeline branch. The model is suitable for noisy data. Multiple leaks can be detected by serial compensation strategy. The method tested on a simulation network indicates that the method is more accurate for single leak case than the multiple one. Experiments carried out using a bench scale laboratory network detected 19 of the 30 simulated leaks.

Cortes and Alejo (1997) conducted extensive field studies to obtain data on real user consumption, volumes lost from house connections, volumes lost from main and secondary lines, meter mis-registry, and unauthorized connections. Stratified random sampling theory is used to determine the size of sample studies, based on initial pressure and leak frequency estimates. The initial estimate is progressively modified to get a better estimate of the sample size. A two-sample study is conducted to determine the percentage of household connections with leaks and estimate the volume of water lost from the house connections. Different approaches have been used to evaluate water losses from main and secondary lines, over and under-registered connections, unauthorized use, and billing measurement errors. Hydrometric district technique (similar to district metering approach) is used to isolate a part of the distribution network from a hydraulic point of view. This method gives accurate estimates of water losses and also the cause for water losses. The method has been applied to 15 Mexican cities.

Vitkovsky, Simpson, and Lambert (1999) developed a unique method to estimate magnitude of leaks using transients and genetic algorithms. Under unsteady conditions of flow, numerically modeled hydraulic grade lines are fitted against measured HGLs. The negative sum of absolute differences between measured and modeled pressure heads are then maximized to solve for Darcy-Weisbach pipe friction factors (f) and lumped leak coefficients. Leakage is simulated using an orifice equation. An average error of 3.43% for the estimated friction factor and 0.5% for the leak parameter is found from the test simulations. However, the method is applied only to a test network consisting of 11 pipes and 7 nodes. The applicability of the method to large pipe networks is not discussed.

The implicit state-estimation technique adopted by Andersen and Powell (2000) detects leaks in district metering areas (DMAs). Pressure and flow measurements are taken at nodal points across a distribution network. A leak is suspected from increased demands or night flows at the nodes. Measurements are simulated in idealized noise-free conditions. Weights are assigned for pressure and flow measurements. When the demand factors for the leaks at nodes are increased, the magnitude of leakage becomes close to the actual demand at the node and hence, leak detection becomes difficult.

Mpesha, Gassman, and Chaudhry (2001) applied the frequency response method to detect and locate leaks in pipelines. The method requires pressure measurements at the valve. A frequency response diagram is created at the valve by periodic opening and closing of the valve. The steady–oscillatory flow thus created at the valve is analyzed in the frequency domain using the transfer matrix method. Resonant pressure amplitude peaks in the frequency response diagrams are compared in the presence and absence of leaks to detect leaks. Extensive measurements of hydraulic parameters are not required in this method and leaks up to 0.5% of the mean discharge can be detected for various values of friction factors. The authors use a similar method to detect leaks in transient flow conditions (Mpesha, Gassman, and Chaudhry, 2002). Transients created by opening and closing of a valve are analyzed using the method of characteristics. The method detects and locates leaks in pipelines with friction factors ranging from 0.01 to 0.025.

Summary of the various leak detection techniques, their usefulness under different conditions, and leak information obtained from each of the methods are given in Table 2.1.

Table 2.1 Overview of some recent leak detection studies

Author and Year	Data Needed	Flow Condition	Mathematical Approach	Leak Information	Network Calibration Requirements	Comment
Pudar and Ligget, 1992	Flow and pressure	Steady	Least squares minimization	Location and magnitude	Required	Requires accurate friction factors
Ligget and Chen, 1994	Flow	Unsteady	Inverse transient analysis by MOC	Location and magnitude	Required	Simultaneous network calibration
Liou and Tian 1995	Pressure	Unsteady	Cauchy algorithm and time-marching algorithm	Location	Not required	Developed for oil pipelines
Mukherjee and Narasimhan, 1996	Flow and pressure	Steady	Likelihood ratio method	Location and magnitude	Required	Detects multiple leaks
Arreguin-Cortes and Ochoa-Alejo, 1997	Flows	Unsteady	Stratified random sampling	Magnitude	Required	Applied to 15 Mexican cities
Liou, 1998	Pressure	Unsteady	Impulse response extraction	Location	Not required	Developed for oil pipelines
Vitkovsky, Simpson and Lambert, 1999	Pressure	Unsteady	Genetic algorithm optimization	Magnitude	Required	Slow convergence
Andersen and Powell, 2000	Flow and pressure	Unsteady	Weighted least squares estimation	Location	Not required	Applied to DMAs
Mpesha, Gassman, and Chaudhry,2001	Pressures at pipe valves	Steady oscillatory flow	Frequency response method	Location and magnitude	Required	Designed for open loop
Mpesha, Gassman, and Chaudhry, 2002	Flow and pressure	Unsteady	MOC with Fast Fourier transform	Location and magnitude	Required	Requires accurate friction factors
Buchberger and Nadimpalli, 2003	Flow	Unsteady	Sequential time-based flow statistics	Magnitude	Not required	Quick and accurate

CHAPTER 3

PIPE FLOW IN RESIDENTIAL SERVICE ZONES

3.1 RESIDENTIAL SERVICE ZONES

The reliability of hydraulic models depends on the accuracy of field measurements. For the leak detection method proposed in the present research, accurate continuous pipe flow measurements are required from a supply line feeding a well-defined residential service zone of a water distribution system. The residential service zones are either branched pipes conveying water to homes in the service area or well-established demand monitoring areas (DMAs, also known as district metering areas in European countries).

In recent years, water distribution systems have been broken down into smaller demand monitoring areas for their efficient management. Special district meters are installed at the main feeder lines of a DMA and water demands within the area are continuously monitored. One of the primary reasons for establishing DMAs in distribution networks is efficient leakage control. To enhance leak detection, continuously measured water demands are transmitted to an IT system through Supervisory Control and Data Acquisition (SCADA) or Telemetry systems. The measured flows are then analyzed to estimate and locate leaks in the distribution system (Andersen and Powell, 2000 (Table 2.1)).

3.2 DESCRIPTION OF FLOW DATA

Flow in the mainline of a DMA is always uni-directional and hence, flow rate cannot be negative. In some network loops, flow can reverse direction and thus can take negative values. Zero flow in a DMA supply line implies stagnant conditions (idle conditions). For the proposed leak estimation method, flows have to be either positive or zero.

It would be ideal to demonstrate the proposed leak detection method using real water demand data with leakage present in them. However, such data were not available for considerably large residential area. Hence, simulated water demand data with artificial leakage rates are used to explain the leak detection method. The three time series plots in Figure 3.1 are integrated pipe flows from a residential service zone consisting of 200 homes. Flows are simulated every second for one hour during low water use period. Water use at individual residences is assumed to follow a Poisson Rectangular Pulse (PRP) process (Buchberger and Wu 1995). The PRP parameter values (Table 3.1) to simulate water demands are taken from recent field studies on residential water use (Buchberger *et al.,* 2003).

Table 3.1 Parameters for Poisson Rectangular Pulse simulation of residential indoor water demands during low use period.

Mean Arrival Rate, λ	0.008 per minute per residence	
Pulse Duration Lognormal Distribution	Mean Variance	8.5 L/min 22.2 $(L/min)^2$
Pulse Duration Lognormal Distribution	Mean Variance	0.75 min 5.70 $(min)^2$

Figure 3.1a shows a flow hydrograph of a main supply line for a one-hour period. This is an ideal condition when there is no leakage in pipe network. The average flow during this one-hour period is 10.46 L/min and the maximum flow approaches 51.2 L/min. Stagnation occurs nearly 30% of the time, as the flow in the supply line drops to zero on several occasions during this period of low demand. The remaining two plots in Figure 3.1 depict the same neighborhood, but show hydrographs with network leakage. In Figure 3.1b, leakage occurs at a constant rate Q_L = 3.79 L/min (1 gpm) and in Figure 3.1c, the leakage rate is variable with mean 9.07 L/min and standard deviation 1.79 L/min. The maximum and minimum leakage rates for the variable leakage case are 15.66 L/min and 1.66 L/min, respectively. It can be observed that stagnation disappears in the presence of pipe leakage.

Analysis of recent field measurements suggests that flow rates through branching lines in residential zones can be well described with a mixed truncated normal distribution (Buchberger *et al* 2003). Consider Figure 3.2 consisting of three normal probability plots for flow sequences described in Figure 3.1. The vertical mass spike at zero flow (Q = 0) of Figure 3.2a corresponds to a stagnation of 29.4 %. The linear part of the plot over Q > 0 indicates a continuous probability distribution of pipe flows when there is water use.

Figure 3.2b shows the case where the pipe network has a constant leakage rate of Q_L = 3.79 L/min (1 gpm) at all times. The probability plot is identical to Figure 3.2a, except that all points have been shifted to the right an amount equal to the constant leakage rate. In the presence of pipe leakage, stagnation is not possible and the minimum flow increases from 0 to 3.79 L/min. The vertical mass spike now centered up on Q = 3.79 L/min is due to pipe leakage only. This is

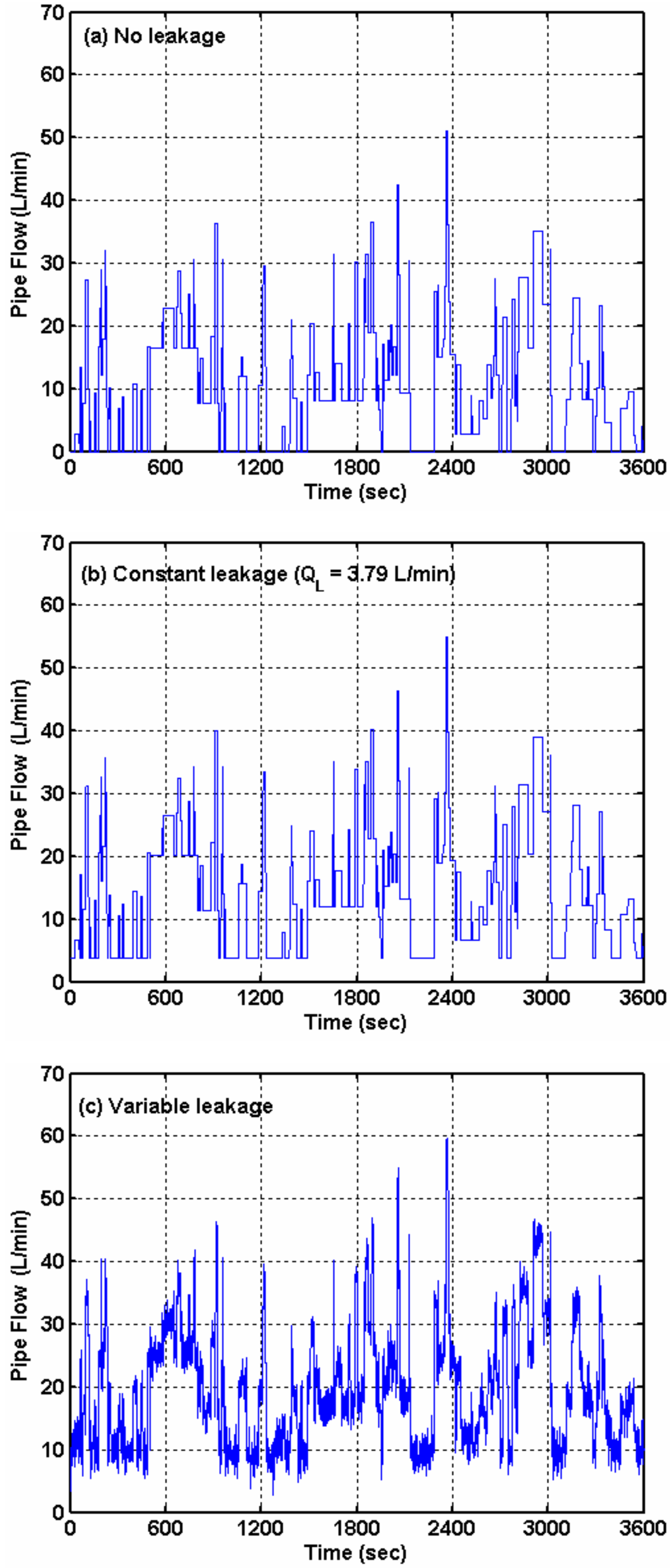


Figure 3.1 Time series plot of simulated flows in a mainline feeding a DMA with 200 homes during a period of low water use

an ideal condition where pipe leakage rate can be estimated from the shift in the vertical spike from Q = 0 to Q > 0.

However, in reality, fluctuations in pipe networks are inevitable due to changes in pump operating conditions, opening and closing of valves, etc. These fluctuations cause statistical noise in the measured pipe flows, and unsteady leakage rates. Figure 3.2c shows the case where the supply line is subject to unsteady leakage. Leakage rate can no longer be predicted from non–uniform displacement of the vertical mass spike. The method proposed in this research is developed to predict leakage rate in the unsteady conditions.

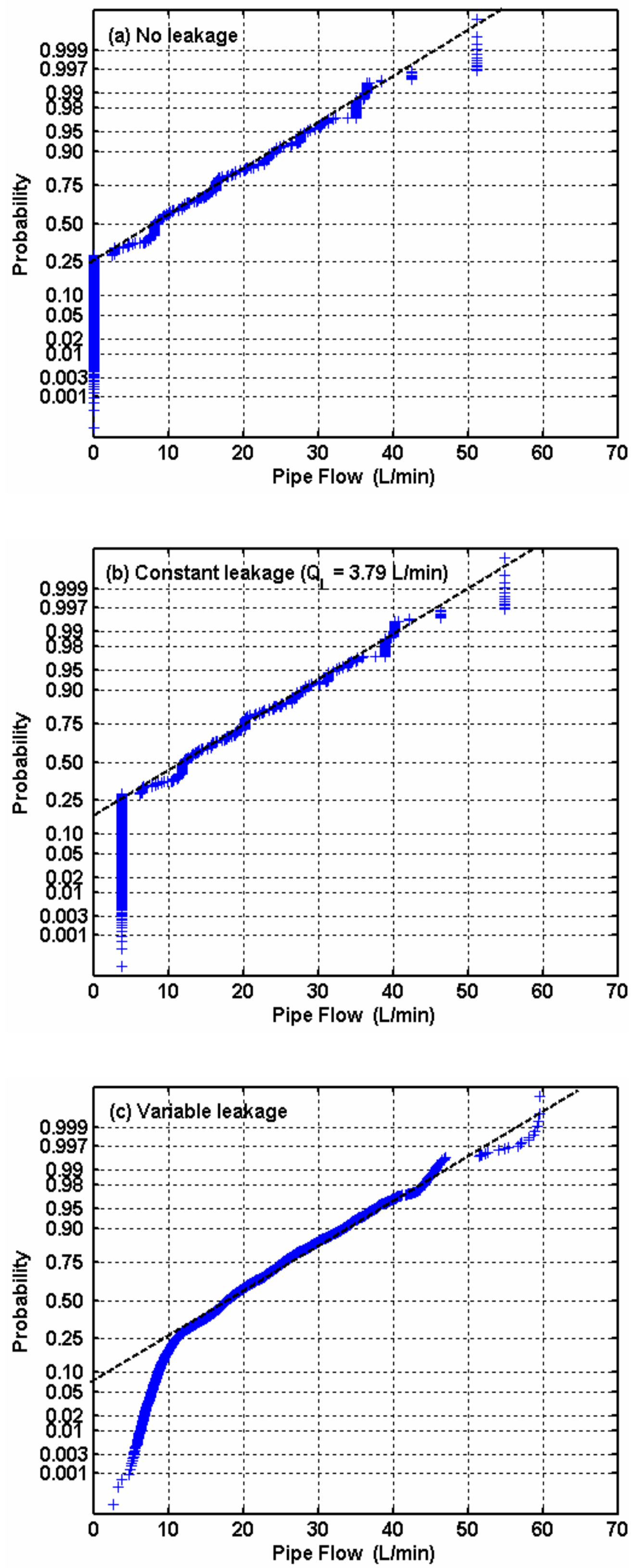


Figure 3.2 Normal probability plots of pipe flow data shown in Figure 3.1

CHAPTER 4

LEAK ESTIMATION METHOD

4.1 MOMENTS FOR MIXED TRUNCATED NORMAL DISTRIBUTION

The rate of water flowing through any pipe in a distribution system can be viewed as a random variable. The normal probability plots shown in Figure 3.2 indicate that the pipe flows can be approximated as a mixed truncated normal distribution. If the upper tail of the flow data does not plot as a straight line on normal probability paper, then a Box-Cox transformation (Hipel and McLeod, 1994) can be applied to all the flow data with the objective of normalizing the upper tail region.

Consider the case with a constant leakage rate Q_L. The mixed truncated normal distribution has two parts:

(a) A discrete probability mass at $Q = Q_L$ and

(b) A continuous probability distribution restricted to the positive number line.

The corresponding moments of the mixed truncated normal distribution are given by

$$E[Q^k] = Q_L^k.\Phi\left(\frac{Q_L - \mu}{\sigma}\right) + \int_{Q_L}^{\infty} \frac{q^k}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right] dq \qquad (4.1)$$

where μ and σ^2 are the mean and standard deviation of the "parent" normal distribution indicated by the linear part of the normal probability plots. The "parent" normal distribution best fits the upper tail of the measured flow data. For example, the dashed lines on the probability plots in Figure 3.2 show the parent normal distribution for each sample. Estimating sample values of μ and σ^2 for the parent distribution from measured flow data is a key step in the proposed leak detection method. For k = 1 and 2, Eq (4.1) gives

$$E[Q] = Q_L.\Phi\left(\frac{Q_L-\mu}{\sigma}\right) + \int_{Q_L}^{\infty} \frac{q}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right] dq \qquad (4.2a)$$

$$E[Q^2] = Q_L^2.\Phi\left(\frac{Q_L-\mu}{\sigma}\right) + \int_{Q_L}^{\infty} \frac{q^2}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right] dq \qquad (4.2b)$$

Expanding the integral parts of the above two equations gives

$$E[Q] = Q_L\Phi\left(\frac{Q_L-\mu}{\sigma}\right) + \sigma\left\{\phi\left(\frac{Q_L-\mu}{\sigma}\right) + \left(\frac{\mu}{\sigma}\right)\left[1-\Phi\left(\frac{Q_L-\mu}{\sigma}\right)\right]\right\} \qquad (4.3a)$$

$$E\left[Q^2\right] = Q_L^2\Phi\left(\frac{Q_L-\mu}{\sigma}\right) + \sigma^2\left\{\left(\frac{Q_L+\mu}{\sigma}\right)\phi\left(\frac{Q_L-\mu}{\sigma}\right) + \left(1+\frac{\mu^2}{\sigma^2}\right)\left[1-\Phi\left(\frac{Q_L-\mu}{\sigma}\right)\right]\right\} \qquad (4.3b)$$

Expressing Equations 4.3a and 4.3b in terms of the standardized truncation level,

$$\varepsilon = \frac{Q_L-\mu}{\sigma} \qquad (4.4a)$$

and rescaling flow rate,

$$Q^* = \frac{Q-Q_L}{\sigma} \qquad (4.4b)$$

the moments of mixed truncated normal distributions can be expressed as

$$E\left[Q^*\right] = \phi(\varepsilon) - \varepsilon\Phi(-\varepsilon) \qquad (4.5a)$$

$$Var\left[Q^*\right] = \left[1 + 2\varepsilon\phi(\varepsilon) + \varepsilon^2\Phi(\varepsilon)\right]\Phi(-\varepsilon) - \left[\varepsilon + \phi(\varepsilon)\right]\phi(\varepsilon) \qquad (4.5b)$$

Derivations for Equations 4.5a and 4.5b are given in Appendix A.

Standardized graphs of $E[Q^*]$ and $S[Q^*] = Var[Q^*]^{1/2}$ are obtained by plotting mean and standard deviation against the truncation level ε (Figure 4.1a). They provide a diagnostic tool in identifying the leakage rate hidden in the measured flow data. Polynomial approximations (Abramowitz and Stegun, 1972) have been used to numerically compute the pdf and cdf terms, $\varphi(\varepsilon)$ and $\Phi(\varepsilon)$. Computations for pdf, cdf, $E[Q^*]$, and $S[Q^*]$ have been done in MATLAB. The codes are provided for reference in Appendix D of the thesis.

As discussed in Chapter 3, fluctuations in the distribution system cause random noise in the measured flow data. In the presence of unsteady leakage rates, the standardized curves of $E[Q^*]$, and $S[Q^*]$ may not be very helpful in estimating leakage rates in pipe flows. Hence, the derivatives of $E[Q^*]$ and $S[Q^*]$ are plotted against the standardized truncation level ε (Figure 4.1b). Equations for derivatives of moments (detailed derivations are provided in Appendix A) are given by

$$E'[Q^*] = -\Phi(-\varepsilon) \tag{4.6a}$$

$$S'[Q^*] = -\frac{\Phi(\varepsilon)}{\Theta(\varepsilon)} \tag{4.6b}$$

where $\Theta(\varepsilon) = \frac{S[Q^*]}{E[Q^*]}$ is the coefficient of variation of the rescaled flow rates.

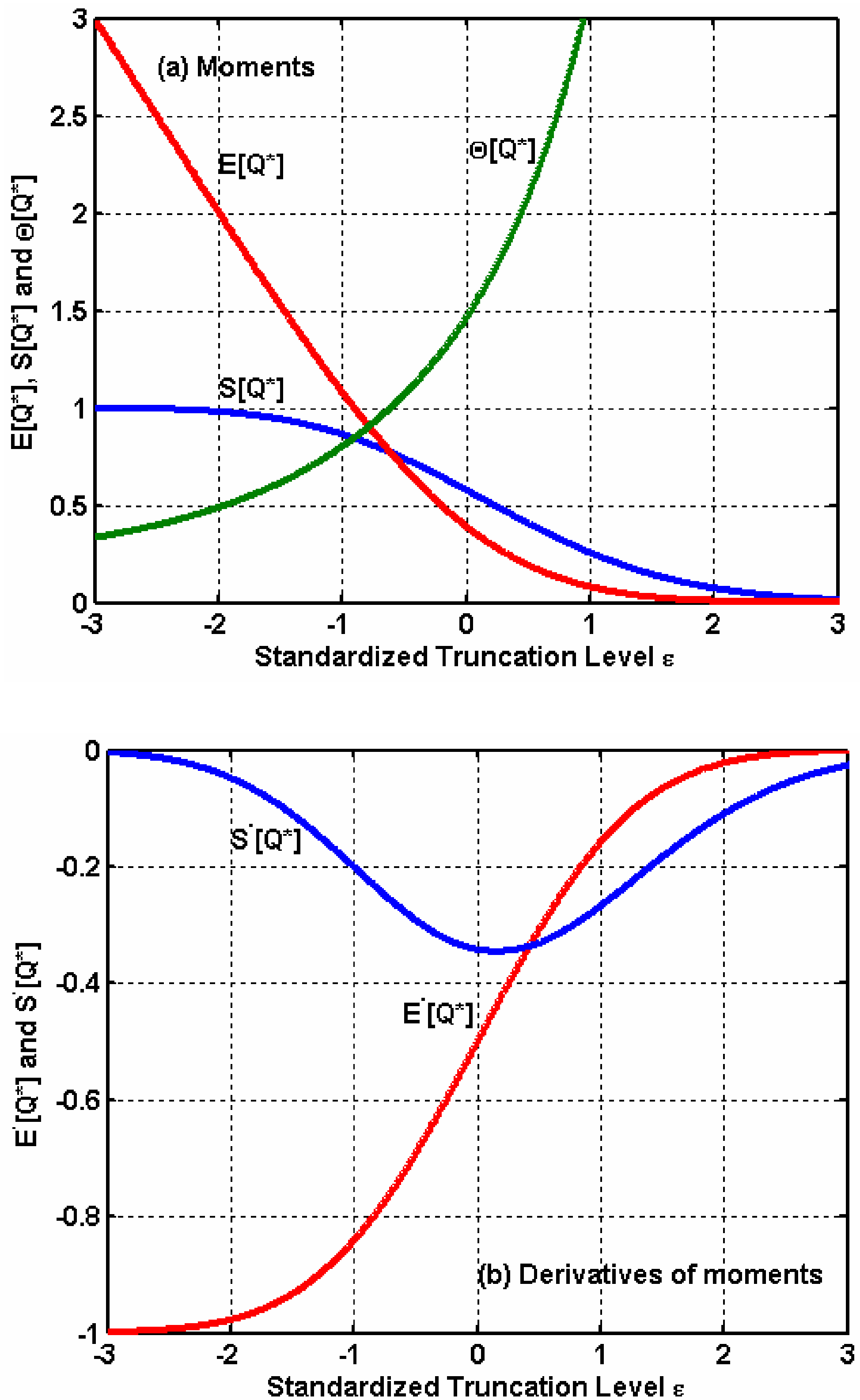


Figure 4.1 Standardized curves for mixed truncated normal distribution

4.2 LEAK DETECTION ALGORITHM

Suppose continuous measurements of flow in a branching pipe are available. Let Q(t) denote the flow at time t and let Q_{MAX} be the maximum value recorded during the observation period. These data can be obtained from network pipes using ultrasonic flow meters with a sampling frequency of one reading per second, preferably taken during a period of low use. For purposes of network leak detection, the ideal time to monitor pipe flows would be for several hours during a winter night with minimum air temperatures above freezing. It is advisable to repeat flow measurements for several nights.

The proposed leak detection method has five steps outlined in Figure 4.2. The basic idea is to repeatedly compute the mean and standard deviation of the measured flows as the data set is progressively truncated from below. The truncation process continues until the entire set of flow data has been reduced to zero. The resulting sample statistics and their sample derivatives are then plotted against the truncation level using the standardized graphs shown in Figure 4.1. The point where the sample values (moments and/or derivatives) diverge from the standardized curve indicates where the truncation threshold matches the maximum network leakage rate. An example will demonstrate this in Chapter 5.

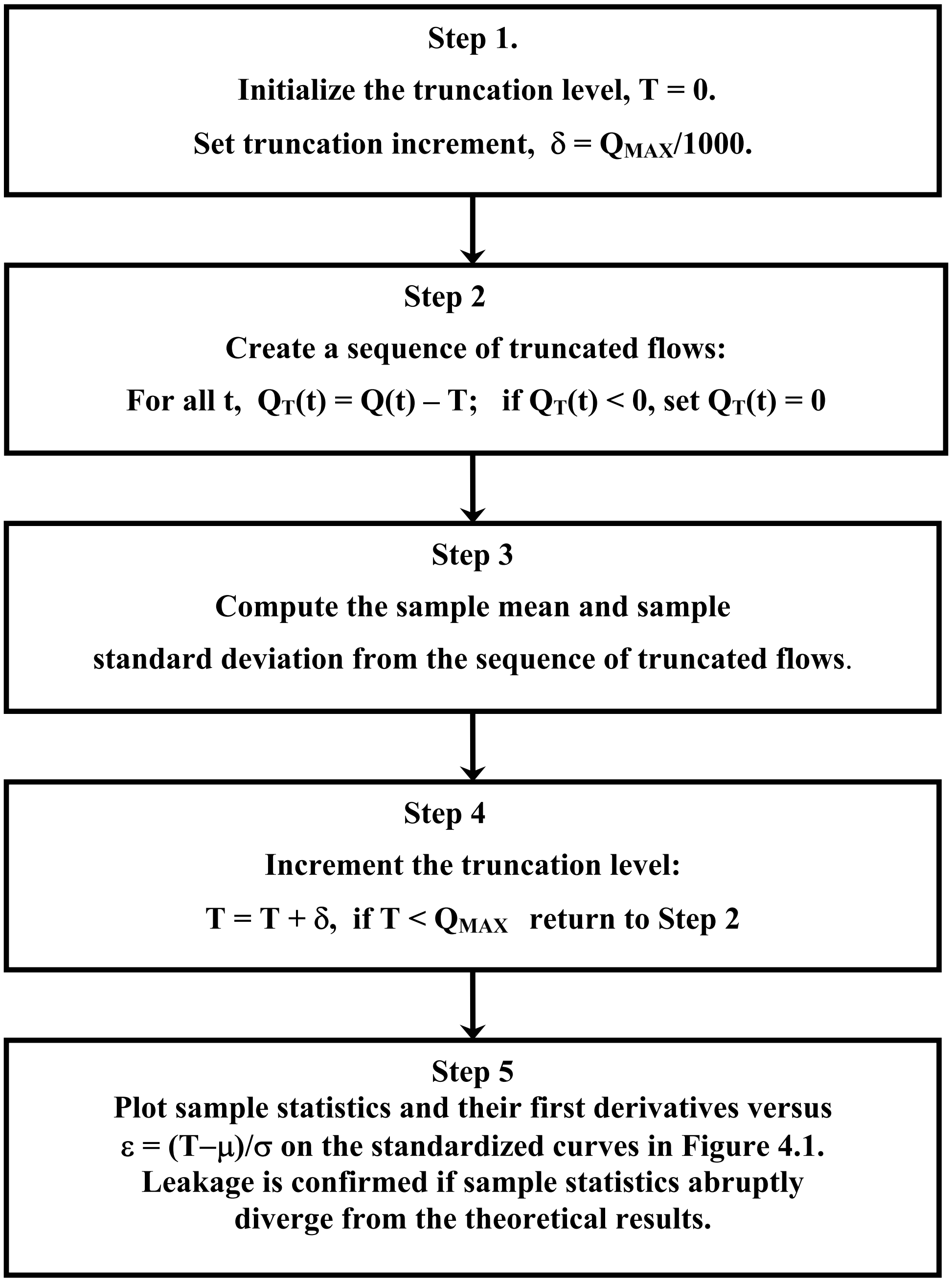


Figure 4.2 Leak detection algorithm

CHAPTER 5

APPLICATION OF LEAK ESTIMATION METHOD

5.1. DATA SOURCES

The leak detection algorithm has been applied to water demand data obtained from two sources.

5.1.1. Milford Data

Residential water demands were collected from a neighborhood of 21 homes in Milford, Ohio for a period of seven months (Buchberger *et al,* 2003.). Flow meters were installed on service lines of all the 21 residences, and water demands were recorded continuously, every second during the entire study period. These water demand data were integrated to obtain continuous pipe flows in the main supply line to the neighborhood. A typical dead-end loop and direction of flows that could occur at the entrance of the loop are shown in Fig. 5.1.

5.1.2. Simulated Data

The second set of data consists of simulated water demands for considerably large neighborhoods. PRP Simulator was used to generate water demands continuously for every second. The PRP parameters (Table 3.1) were taken from Milford field studies discussed in the section 5.1.1.

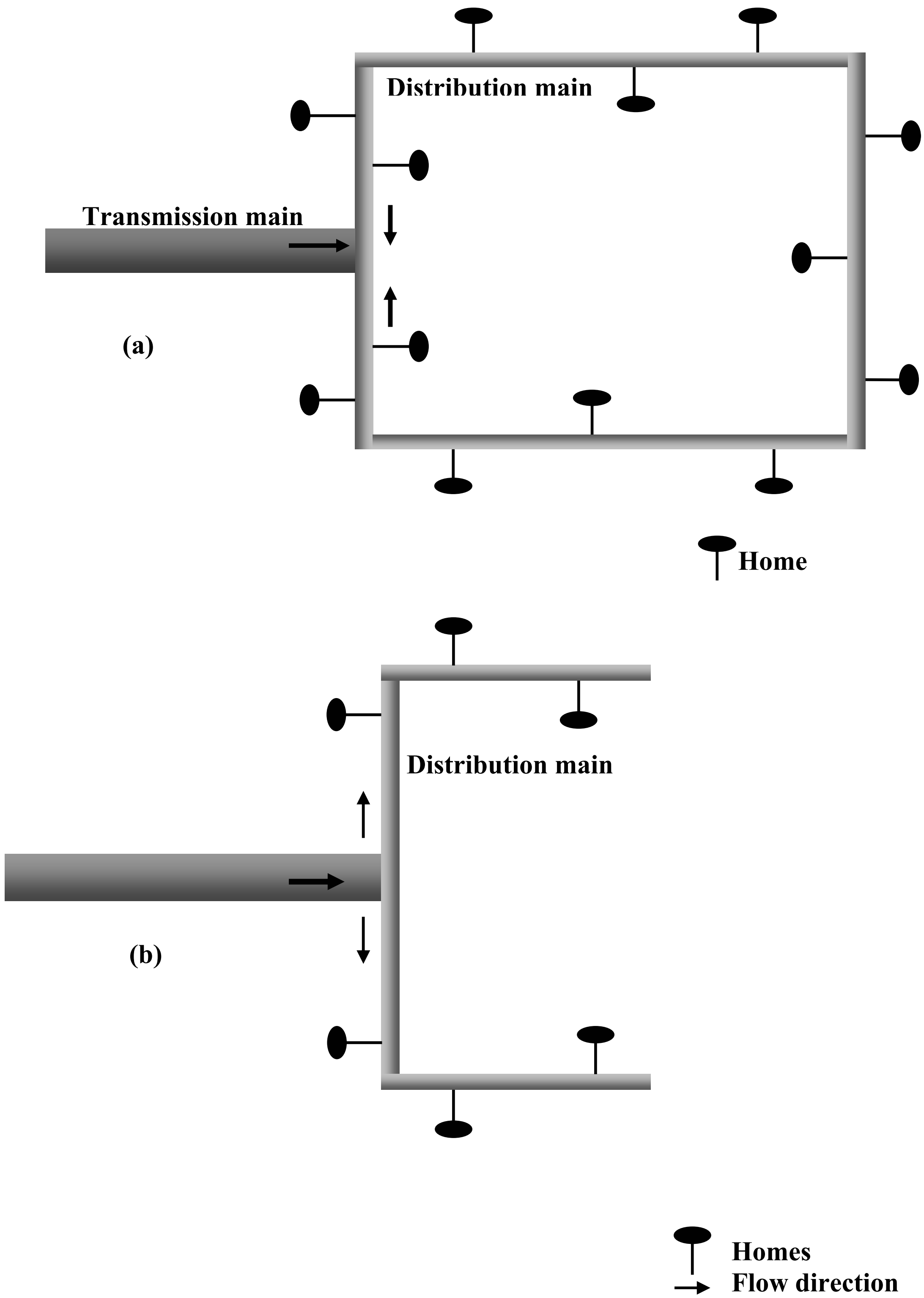


Figure 5.1 (a) Dead–end loop (b) Dead-end stem

5.2. SIMULATION OF LEAKAGE RATES

Both data sets used in this study do not have any leakage in them. Hence, artificial leakage rates were added to residential water demands. In one case, a constant leakage rate of 3.79 L/min was added to all the flows. This is an example for the constant leakage rate shown in Figures 3.1b and 3.2b. In the second case, unsteady leakage rates ranging from 2 L/min to 16 L/min were added to the flow sequences. The leakage rates were generated with a lag-1 auto regressive model, given by Eq (5.1).

$$Q_L(t) = \mu_L(t) + r(1)\left[Q_L(t-1) - \mu_L(t)\right] + \sigma_L \sqrt{\left[1 - r(1)^2\right]}\, Z(t) \tag{5.1}$$

where $r(1)$ = lag-1 autocorrelation coefficient for the leaks

Z = ~N(0,1) is a standard normal variate

σ_L = standard deviation of the leaks and

$\mu_L(t)$ = mean leakage rate at time t

$$\mu_L(t) = Q_L^{\max} - bQ(t) \tag{5.2}$$

In Eq (5.2), $Q(t)$ = pipe flow rate at time t

$Q_L^{\max}$ = maximum mean leakage rate, and

b ≥ 0 , dimensionless constant.

The auto regressive model used to generate random leakage rates is arbitrary; however, it adds statistical noise to the mean leakage rates within certain reasonable bounds. The variable b is assumed to be greater than zero with the assumption that pressures in pipes decrease with increasing flows. The value of leakage parameter r is arbitrary and mimics the pressure fluctuations in a distribution system in Dover Township area, New Jersey (Maslia *et al.*, 2000).

Figure 5.2a shows the simulated instantaneous leakage rates $Q_L(t)$ versus the pipe flow rates $Q(t)$ based on

$r(1) = 0.50$

$\sigma = 0.20\mu_L(\mathrm{t})$

$b = 0.04$

$Q_L^{\max} = 9.5$ L/min

The corresponding time series of simulated leaks, shown in Figure 5.2b, mimics patterns for network pressure fluctuations measured in the field (Maslia *et al*, 2000). The simulated leaks combined with the residential demands provide a realistic picture of continuous pipe flows into a hypothetical neighborhood of 200 homes during a period of low water use (see Figures 3.1b and 3.1c with one value per second for a period of one hour).

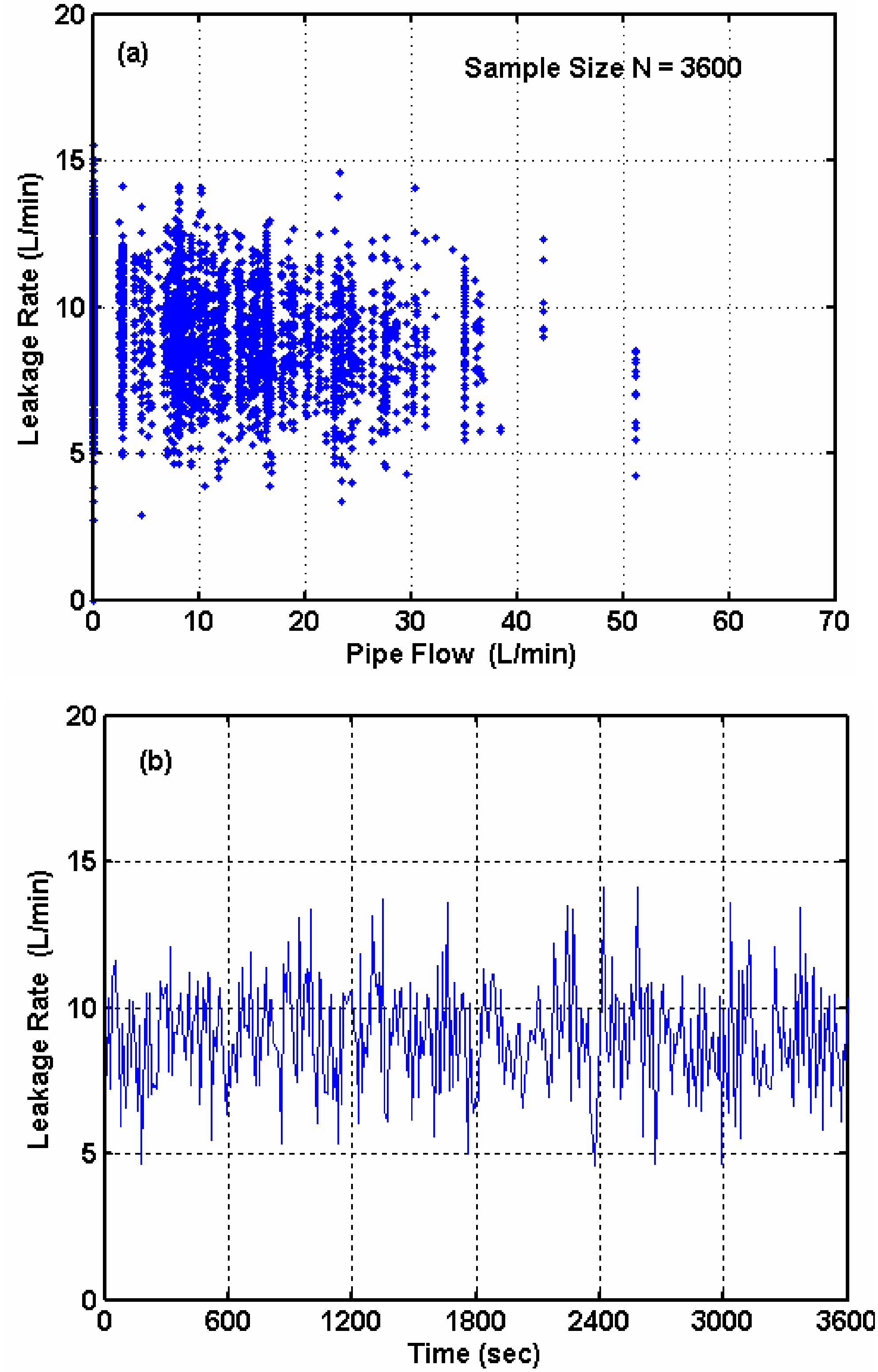


Figure 5.2 Simulated pipe leakage rates versus (a) flow and (b) time.

5.3 ESTIMATING PARENT NORMAL DISTRIBUTION

The positive tail region of the measured flow data of the parent normal distribution is considered to fit a line of form given in Eq. (5.3) using linear regression.

$$Q = \mu + \sigma Z \tag{5.3}$$

In Eq (5.3), μ and σ = parameters of the parent normal distribution, estimated as regression coefficients,

dependent variable *Q* = pipe flow (water demands plus unknown leaks) and

independent variable $Z = \Phi^{-1}(p)$ is the p^{th} quantile of the standard normal distribution found with Blom's plotting position (Stedinger *et al*, 1993),

$$p = 1 - \frac{m - 0.375}{M + 0.250} \tag{5.4}$$

Here *m* is the rank of the observed pipe flows sorted in descending order (maximum value has *m*=1) and M is the total number of flow values in the subset of data points from the positive tail region.

5.4 LEAK DETECTION EXAMPLES

5.4.1 Constant Network Leakage

For the constant leakage example shown in Figures 3.1b and 3.2b, the background rate of network leakage (Q_L=3.79 L/min) is obvious by inspection. However, it would serve as a preview to the variable leak case. In Figure 3.2b, the subset of flow measurements comprising the linear portion of the positive tail is roughly all values where Q > 4 L/min. Using least squares regression, estimates of parameters for the parent normal distribution in Eq (5.1) are μ = 12.68 L/min and σ = 12.09 L/min, with correlation coefficient R = 0.99. This parent distribution appears as the straight line in Figure 3.2b.

Following the 5-step algorithm outlined in Figure 4.2 (with Q_{MAX} = 54.95 L/min), the non-dimensional sample mean and standard deviation for each complete sequence of truncated flows $Q_T(t)$ are computed using Eqs (5.5a) and (5.5b), respectively,

$$\overline{Q^*} = \frac{1}{N\sigma}\sum_{t=1}^{N} Q_T(t) \tag{5.5a}$$

$$S^* = \frac{1}{N\sigma}\sqrt{N\left(\sum_{t=1}^{N} Q_T^2(t)\right) - \left(\sum_{t=1}^{N} Q_T(t)\right)^2} \tag{5.5b}$$

The sample statistics are plotted against the standardized truncation level, $\varepsilon = (T-\mu)/\sigma$, as shown in Figures 5.3a and 5.3b. The standardized curves correspond to the condition $Q_L = 0$. Hence, the deviations between curves of computed sample statistics and those of standardized moments indicate presence of network leakage. Let ε_L denote the point of departure between the trajectories for the sample statistics and the standardized curve, as measured along the abscissa. Then, rearranging Eq (4.4a), the magnitude of the leakage rate is given by

$$Q_L = \mu + \sigma\varepsilon_L \tag{5.6}$$

In Figure 5.3, it is observed that the point of departure between the sample statistics and standardized curves occurs at ε_L = -0.72 This value can be read directly from the abscissa on Figure 5.3 and also can be confirmed by examining numerical differences between computed sample statistics and the standardized curves. From Eq (5.6), the estimated rate of network leakage is Q_L = (12.68 L/min) + (12.09 L/min)(-0.72) = 3.98 L/min. This estimate of network leakage agrees closely with the known constant value of Q_L = 3.79 L/min.

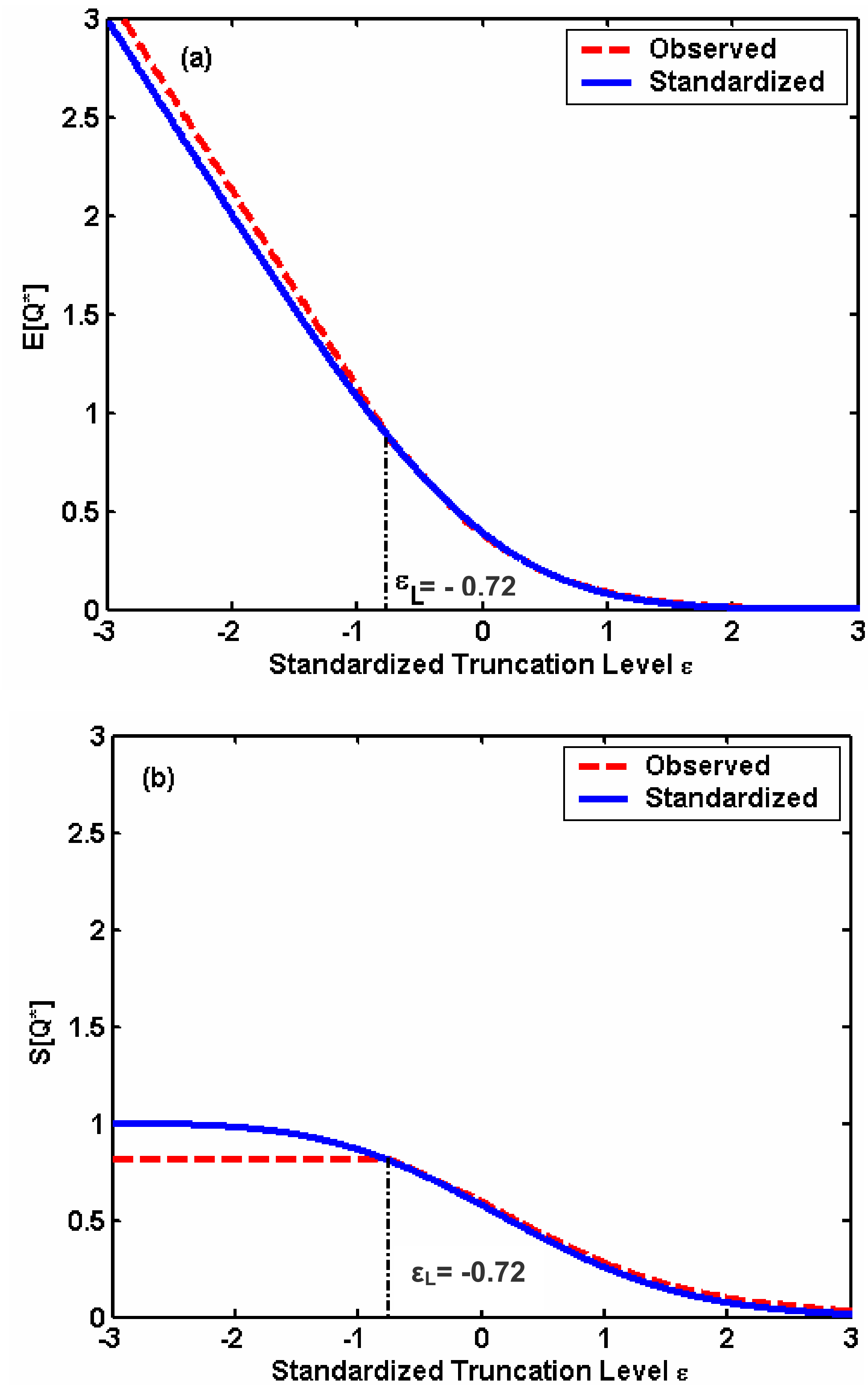


Figure 5.3 Observed and standardized curves for constant leakage example, (a) mean and (b) standard deviation.

5.4.2 Variable Network Leakage

In reality, leakage is not constant as shown in the time series plot and normal probability plots of pipe flows in Figures 3.1c and 3.2c respectively (reasons for unsteadiness are discussed in the previous sections of the chapter). The point of departures between sample trajectories and standardized curves may not be very crisp due to this unsteadiness in the pipe leakage rates. Hence, the 5-step algorithm used for the constant leak example is extended by an additional step, which involves plotting the derivatives of sample statistics and standardized moments. The slopes for the standardized curves are given in Eqs (4.6a) and (4.6b) and plotted in Figure 5.4. Estimates of the slopes for the sample statistics are computed using

$$y'_n = \frac{y\left(T_{n+1}\right) - y\left(T_n\right)}{T_{n+1} - T_n} \qquad (5.7)$$

where y'_n = slope of the sample moment and

$y(T_n)$ = estimated sample moment at the n^{th} truncation level.

In Figure 3.2c, the subset of flows comprising the linear portion of the positive tail corresponds to all values where Q >15 L/min. The least squares estimates for parameters of the parent normal distribution for this tail region are μ = 17.67 L/min and σ = 12.18 L/min, with R=0.99. This parent distribution is plotted as the straight line in Figure 3.2c.

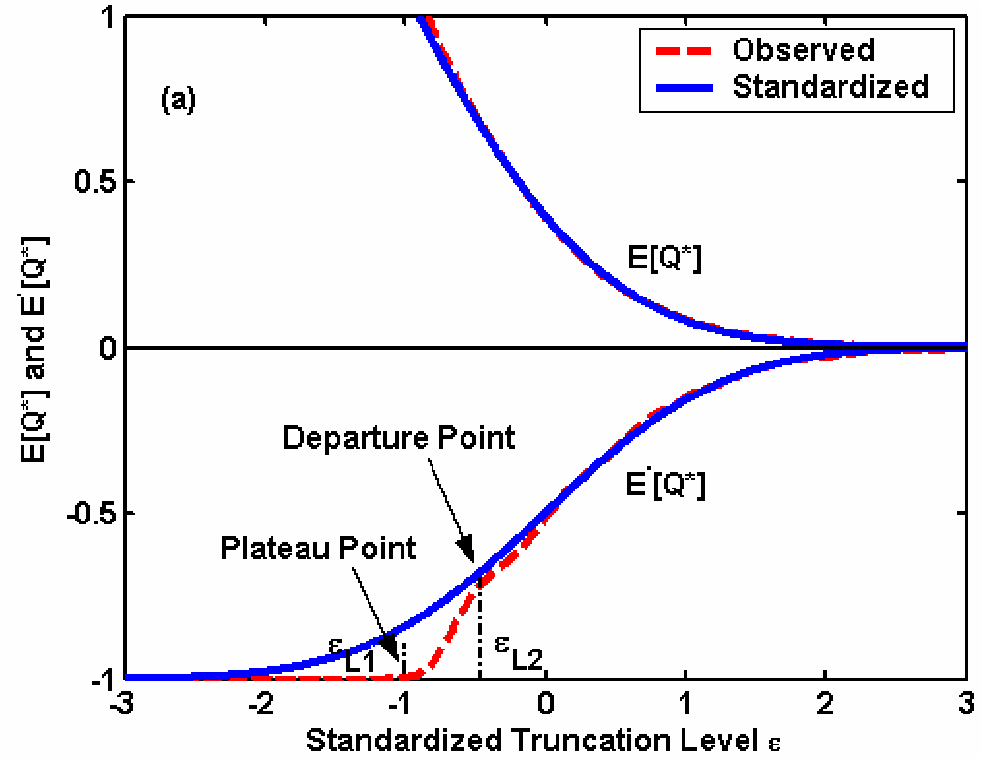
The sample mean and standard deviation for each of the truncated flow sequences are computed using Eqs (5.5a) and (5.5b). The derivatives for sample mean and standard deviation are computed with Eq (5.7). These values along with the standardized curves (given by Eqs (4.5) and (4.6) are then plotted against the standardized truncation level, $\varepsilon = (T-\mu)/\sigma$, in Figures 5.4a and 5.4b.

With the variable leakage case, there are two points of interest on the graphs for the derivatives: (i) the "plateau point", where the sample trajectory flattens and stays horizontal and (ii) the "departure point", where the sample trajectory pulls away from the standardized curve. These points are identified in Figures 5.4a and 5.4b where the corresponding standardized truncation levels, read directly from the abscissa, are estimated to be: ε_{L1} = -1.00 (at the plateau point) and ε_{L2} = -0.45 (at the departure point). Precise estimates of these two points can be obtained by examining numerical differences between the computed slope of the sample statistics and the standardized curves. Both of the standardized truncation levels (i.e., ε_{L1} and ε_{L2}) are used in Eq (5.6) to provide a range of network leakage rates:

Lower Value: Q_L = (17.67 L/min) + (12.18 L/min)(-1.00) = 5.49 L/min.

Upper Value: Q_L = (17.67 L/min) + (12.18 L/min)(-0.45) = 12.19 L/min.

The predicted limits of network leakage do not reach the simulated extremes (2.73 to 15.53 L/min). However they do encompass 94 percent of the simulated leakage values.



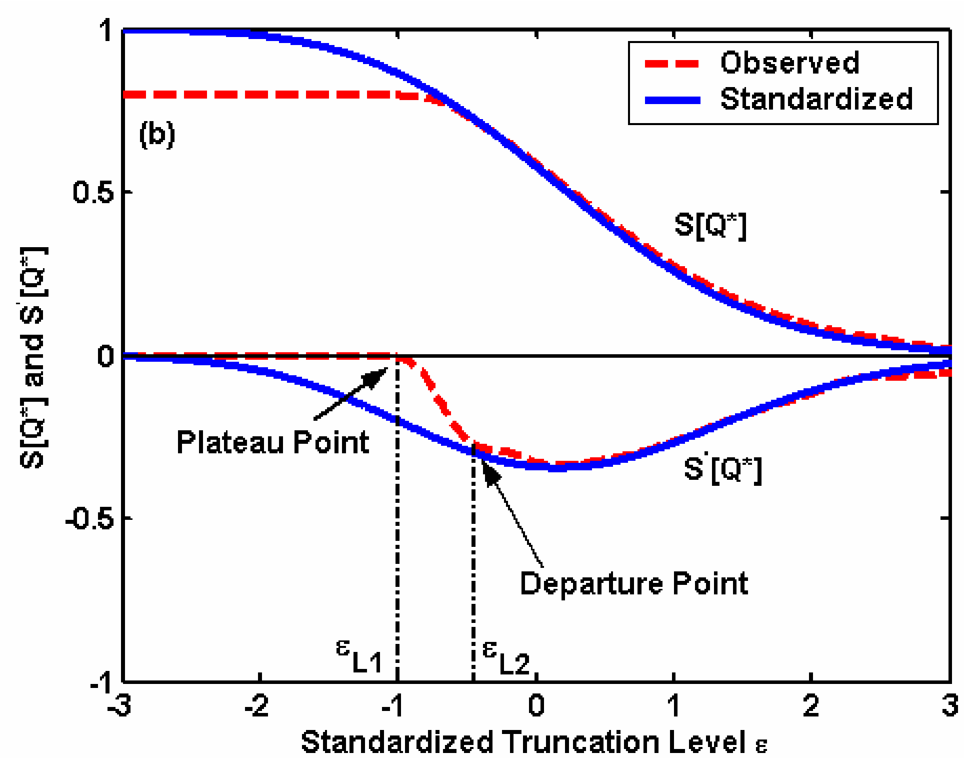


Figure 5.4 Observed and standardized curves for variable leakage example (a) mean and its derivative and (b) standard deviation and its derivative

5.4.3 More Examples

The detection algorithm was tested further against constant and unsteady leakage rates using several other data sets. Fourteen of the tests involved pipe flows constructed from residential water demands measured at 21 homes during a 7-month period extending from April through October. Four other tests involved PRP simulated pipe flows for hypothetical neighborhoods of 400 and 500 homes. Results for the constant leakage case, summarized in Table 5.1, show that the proposed method is quite accurate, usually placing the estimated network leakage rate within 4% of the actual background leakage rate. Results for the variable leakage case, listed in Table 5.2 and depicted on the Box and Whisker plots in Figure 5.5 also show very good agreement between the estimated leakage rates and the simulated leakage rates. While the extreme values are not reached, the predicted range usually covers about 90 percent of the actual rate of network leakage. The plots of computed and standardized moments for the flows for the data sets listed in Tables 5.1 and 5.2 are included in Appendices B1 and B2 respectively.

The leak detection algorithm was also tested on real pipe flow data suspected of leakage. Pipe flows were recorded continuously for three days at a main line feeding an apartment complex of 2000 flats in Poland. The results are discussed in Appendix C of the thesis.

Table 5.1 Simulated and estimated leakage rates for the constant leakage case

Water Demand Data	No. of Homes	Simulated Leakage Rate (L/min)	μ (L/min)	σ (L/min)	ε_L	Estimated Leakage Rate (L/min)	Percent Difference
April	21	3.79	3.61	9.61	0	3.61	-4.8
May	21	3.79	7.35	22.86	-0.15	3.92	3.4
June	21	3.79	1.99	13.43	0.13	3.74	-1.3
July	21	3.79	3.02	10.18	0.09	3.97	4.7
August	21	3.79	2.28	8.75	0.17	3.77	-0.5
September	21	3.79	7.51	11.04	-0.34	3.76	-0.8
October	21	3.79	0.51	14.14	0.23	3.76	-0.8
Simulated	200	3.79	12.68	12.09	-0.72	3.98	5.0
Simulated	400	3.79	14.46	15.8	-0.68	3.72	-1.9
Simulated	500	3.79	18.96	14.94	-1.0	4.02	6.1

Table 5.2 Simulated and estimated leakage rates for variable leakage case

Water Demand Data	No. of homes	Simulated Leakage Rates (L/min)	μ (L/min)	σ (L/min)	ε_{L1}	ε_{L2}	Estimated Leakage Rates (L/min)	Intersection between simulated and estimated leakage rates (%)
April	21	2.16 - 15.06	7.79	9.65	-0.25	0.75	5.37 - 15.03	97.5
May	21	1.14 - 15.25	12.07	21.08	-0.30	0.00	5.75 - 12.07	85.3
June	21	2.46 - 14.99	8.89	11.69	-0.30	0.25	5.41 - 11.77	90.2
July	21	2.38 – 16.35	7.57	10.87	-0.60	-0.10	4.86 - 12.69	96.1
August	21	2.84 - 16.39	9.69	7.34	-0.60	0.25	5.26 - 11.58	87.5
September	21	3.07 - 15.37	13.02	9.77	-0.75	-0.25	5.72 - 10.56	75.8
October	21	2.35 - 15.71	9.50	11.32	-0.40	0.25	4.96 -12.34	94.7
Simulated	200	2.73 - 15.53	17.67	12.18	-1.00	-0.45	5.49 - 12.19	94.2
Simulated	400	3.48 - 15.67	18.91	16.29	-0.84	-0.20	5.22 - 15.66	98.5
Simulated	500	2.46 - 14.84	23.97	14.96	-1.27	-0.75	4.97 - 12.75	97.8

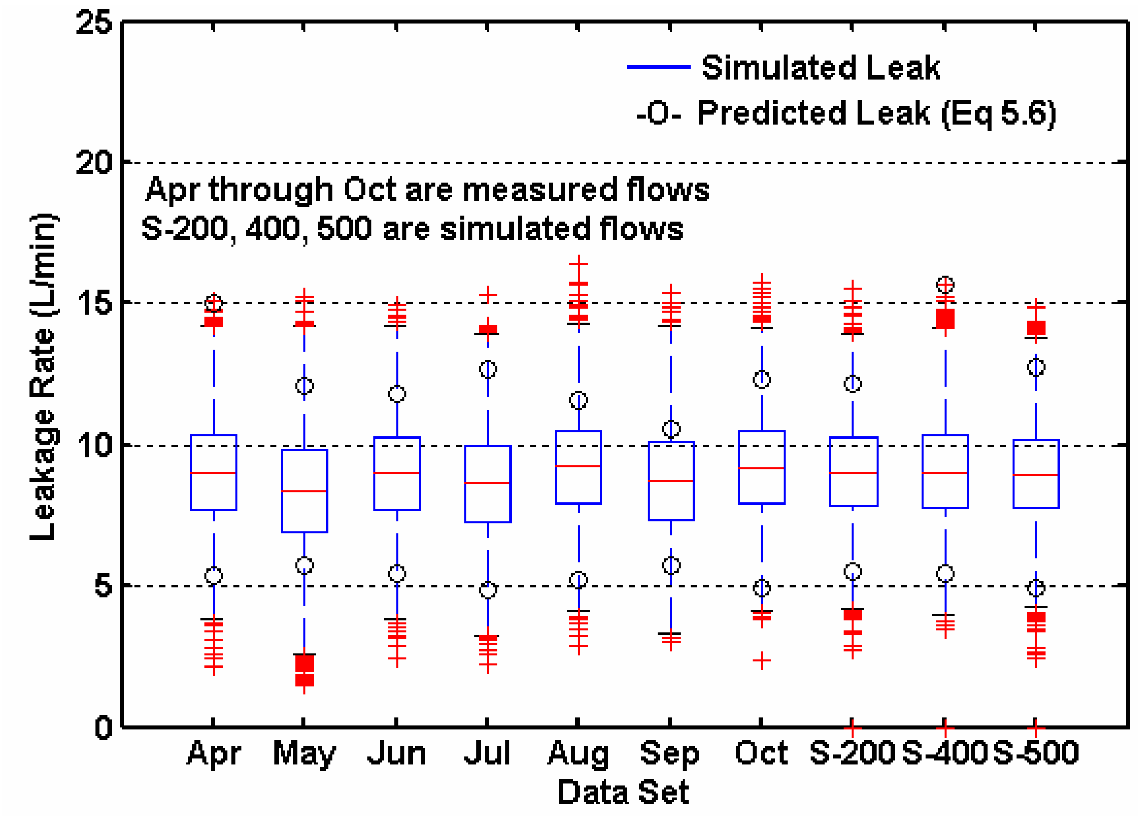


Figure 5.5 Comparison of simulated and predicted leakage rates

CHAPTER 6

PERFORMANCE LIMITS

Performance of the leak detection algorithm depends on the percentage of stagnation in the flows measured at a DMA supply line. In the absence of leakage, stagnation can be identified based on two scaling factors.

1. Spatial scale, in terms of size of neighborhood.
2. Temporal scale, in terms of the averaging time step of pipe flow measurements.

The temporal scale of flow measurements must not be so large that the time-averaged busy periods subsume the idle periods.

Figure 6.1 shows the reduction in stagnation with increasing time step of flow measurements. The time steps are varied from 10 seconds to 300 seconds. Mass is preserved during the flow averaging step, but the averaging reduces the resolution (and hence, the information content) of the flow signal. Performance limits for the leak detection method can thus be defined in terms of size of the maximum service zone and the maximum size of time step.

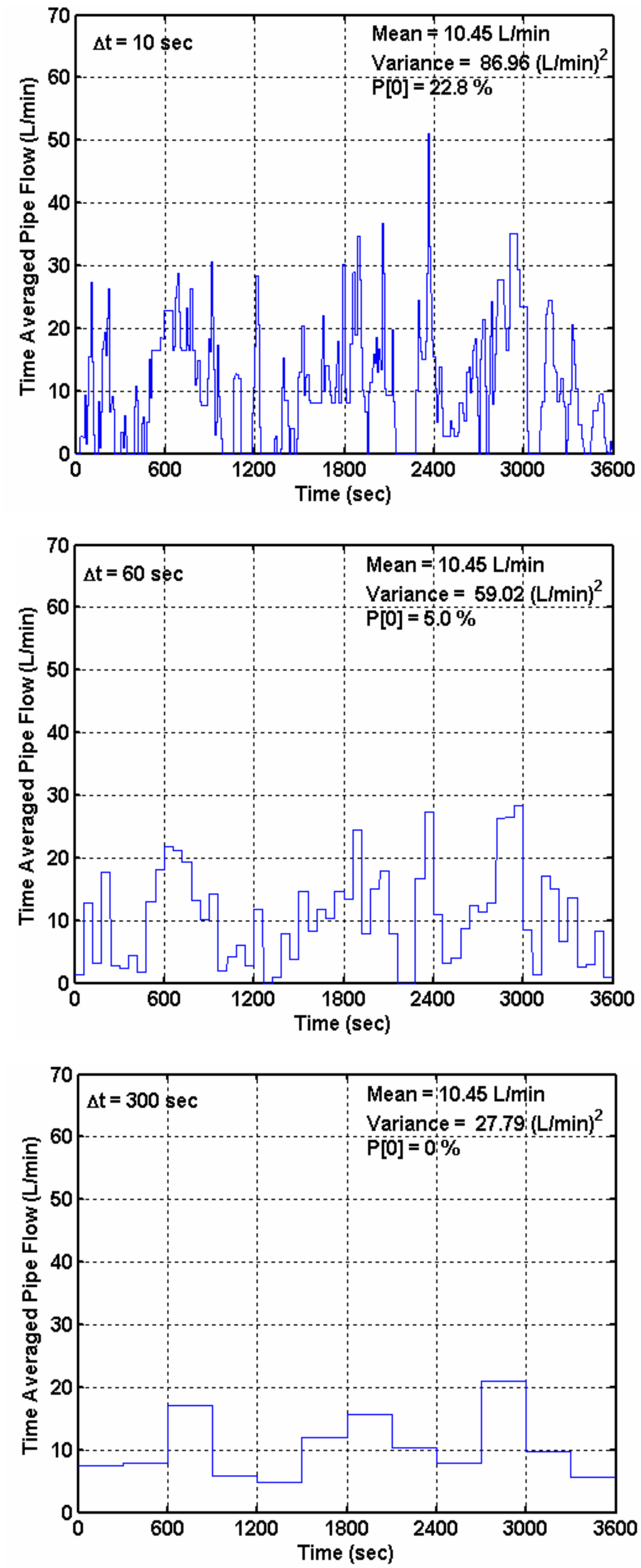


Figure 6.1 Effect of time step on flows measured in a DMA supply line

6.1 PREDICTION OF STAGNTION

Consider the instantaneous rate of flow over a short period (e.g., few hours) through a supply line into a DMA having N homes. Suppose that during this time, residential water use in the DMA follows a homogeneous PRP process with an average arrival rate λ at each household and mean pulse duration τ. The probability of stagnation (i.e., no busy homes in the neighborhood) at any instant in time is given by (Buchberger and Wu, 1995)

$$\mathrm{P}\left[\mathrm{Q{=}0}\right] = \exp(-N\lambda\tau) \tag{6.1}$$

Now, assume that a time averaging grid of Δt is superimposed on the sequence of instantaneous flow readings, and the averaged flow is computed for each time interval Δt. A stagnant moment will remain stagnant only if no PRP arrivals (no water demands) occur during the entire averaging interval. Let K be the number of arrivals during the averaging interval. From the Poisson process, the probability of having no arrivals during the fixed interval Δt is,

$$\mathrm{P}\left[\mathrm{K{=}0}\right] = \exp(-N\lambda\Delta t) \tag{6.2}$$

The presence of stagnation, after a time-averaging operation, implies the joint occurrence of zero flow from the instantaneous process (see Eq (6.1)) and zero arrivals during the fixed time interval (see Eq (6.2)). Because both events are independent, the probability of stagnation after time averaging is given by

$$P[0] = P[Q=0] \cdot P[K=0] = \exp[-N\lambda(\tau + \Delta t)] \quad (6.3)$$

When Δt = 0, Eq (6.3) reduces to P[0] = exp(-Nλτ), the stagnation probability for the instantaneous homogeneous PRP process. When Δt→ ∞, Eq (6.3) reduces to P[0] = 0. Hence, stagnation is completely lost if the time interval for averaging gets too large. This effect is evident in Figure 6.1.

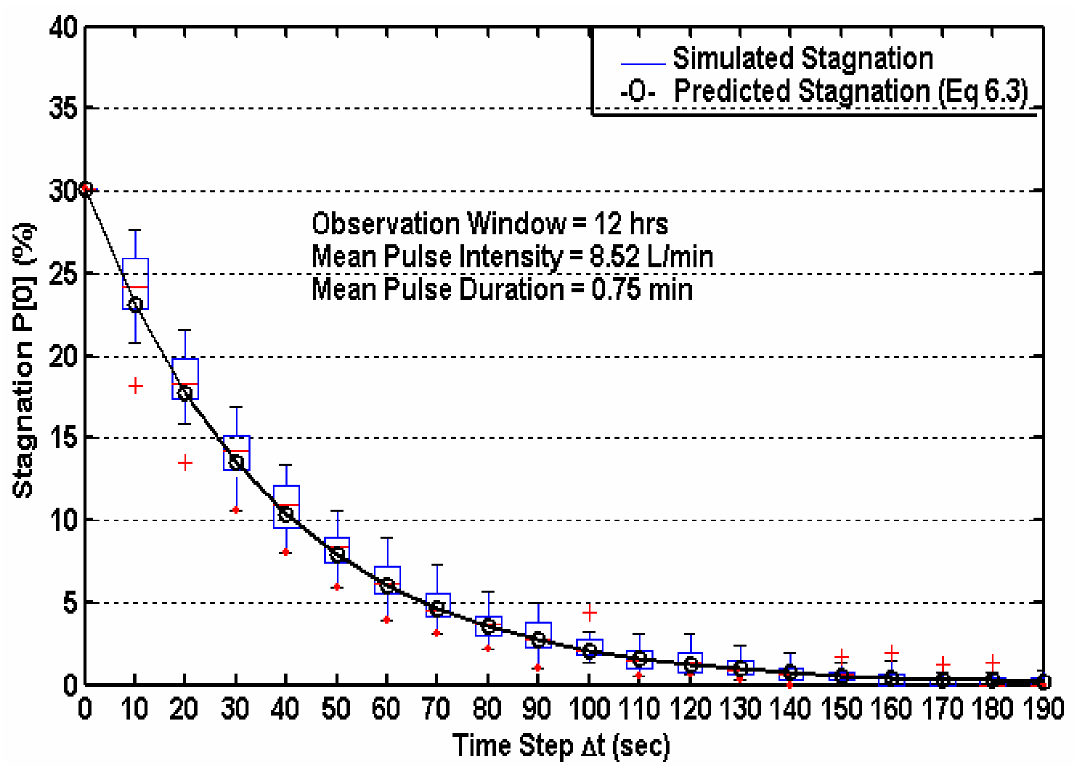


Figure 6.2 Effect of averaging time step on stagnation probability

Stagnation probability decreases as averaging time step increases.

Consider a DMA of 200 homes. PRP simulations were run to generate water demands for the 200 homes with mean arrival rate $\lambda = 0.008$ per minute per home, and mean pulse duration $\tau = 0.75$ min. The demands were integrated to obtain flows through a pipeline feeding the residential DMA. Figure 6.2 illustrates P[0] versus Δt as predicted with Eq (6.3). The stagnation values generated during the PRP simulations are depicted with Box and Whisker plots. There is very good agreement between observed and predicted stagnation probabilities over the full range of time-averaged steps. For the instantaneous flow process ($\Delta t = 0$), Eq (6.1) gives a stagnation probability, P[Q=0] = exp[(-200)(0.008)(0.75)] = 0.30. However, as the averaging time step to record pipe flows is increased to 3 minutes, the stagnation is reduced to zero percent and the hence, the ability of the method to detect leaks in the residential DMA also decreases.

Rearranging Eq (6.3) leads to simple performance limit for the proposed leak detection method,

$$N = \frac{-\ln\left(\mathrm{P}\left[0\right]\right)}{\lambda(\tau + \Delta t)} \tag{6.4}$$

Here performance is expressed in terms of the number of homes that can be screened for leakage in a DMA. Eq (6.4), shows there are four ways to increase performance:

(i) Minimize the resolvable stagnation probability P[0],

(ii) Minimize the mean arrival rate λ ,

(iii) Minimize the mean pulse duration τ ,

(iv) Minimize the averaging time step Δt.

Of these four factors, only the last can be strictly controlled. Nonetheless, thoughtful planning and careful execution can find and exploit near optimal conditions that will help to maximize performance of the leak detection method. For instance, to minimize arrival rates, take flow readings during periods of low demand (i.e., early morning hours). To minimize the mean pulse duration, take flow readings during the winter season, avoiding lengthy outdoor demands. To minimize the averaging time step, increase the frequency of flow recordings (i.e., one every ten seconds or five seconds or finer). In other words, the proposed leak detection method will be most effective if continuous high resolution measurements of pipe flows are obtained during periods of low indoor water use.

The performance limit in Eq (6.4) is plotted in Figure 6.3 for an (early morning) average arrival rate λ=0.005 per minute per home, an average indoor pulse duration τ =1 minute and a series of flow averaging time steps ranging from 1 second to 1 hour. As an example, suppose the minimum resolvable stagnation frequency is P[0] = 10%. When the averaging time step to record flows is 1 second, 470 homes of a DMA can be monitored for leaks at one time. If the averaging time step is now increased to 60 seconds, the maximum size of the neighborhood will reduce to 250. Hence, depending on the size of neighborhood, Figure 6.3 can be effectively used to estimate the length of averaging time step to record flows and improve the performance of the leak detection method.

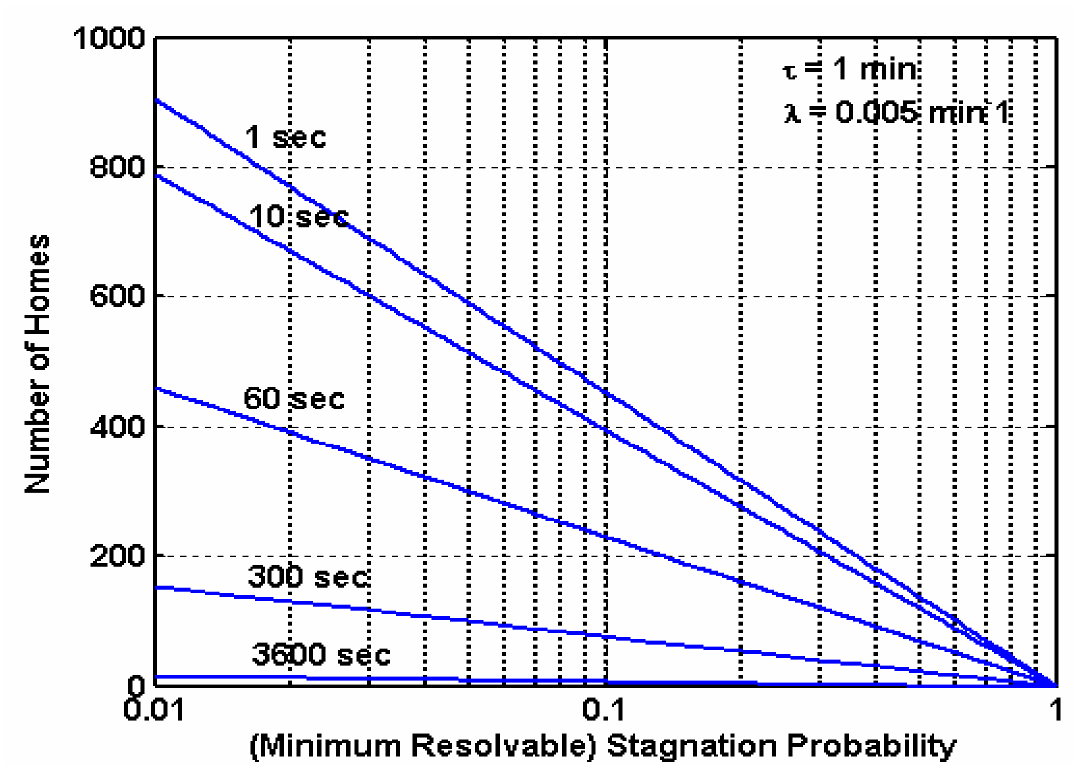


Figure 6.3 Size of DMA vs. stagnation probability

Size of DMA (number of homes) that can be screened for leaks decreases as the size of the averaging time step increases

CHAPTER 7

SUMMARY AND CONCLUSIONS

A new method to estimate magnitude of leaks in pipe networks has been developed and demonstrated. The method is based on sequential statistical analysis of continuous high resolution measurements of discharge taken at one location along a pipe feeding a well defined service zone (e.g., DMA) of up to 500 to 1000 homes. Flows measured from a supply line feeding a DMA are assumed to have a mixed truncated normal distribution, and standardized curves are developed for the mean and standard deviation of flows for an ideal condition of no network leakage. The mean and variance of measured flows are computed as the flows are truncated from the bottom until the truncation level reaches the maximum flow rate within the observation period. The moments of truncated flows are then plotted against the standardized truncation level, and superimposed on the theoretical curves. Presence of leakage in pipe flows is detected if the curves of computed statistics diverge from the standardized curves.

Results obtained indicate that the proposed leak detection method performs well even in the presence of statistical noise in flow measurements. For unsteady pipe flows, maximum and minimum flow rates can be predicted. This method is an improvement over simple water audits currently used by water utilities to estimate leakage rates. The method is also computationally less intensive than existing hydraulic models, and can be applied to steady or unsteady flow conditions.

The performance of the leak detection method can be improved if sufficient stagnation is present in the measured flow readings. As shown in Figure 6.3, a maximum of about 1000 homes can be surveyed for leaks using this approach. This can be achieved by taking flow measurements at night times, when water use is minimum, or during winter times at temperatures above freezing point. It is also important that no negative flows (arising from flow reversals in distribution networks) are recorded. The recorded flow data have to be normal as the method is based on the assumption that flows are normally distributed. Care must be taken to exclude outdoor water use in flow measurements, as the outdoor water use is usually accompanied by long hours of water usage and hence, confound the actual leaks in the flows.

REFERENCES

Abramowitz, M. and I.A. Stegun (1972) **Handbook of Mathematical Functions,** National Bureau of Standards Applied Mathematics Series 55, US Government Prining Office, 1046 pgs.

Arreguín-Cortes, F.I. and L.H. Ochoa-Alejo (1997) "Evaluation of Water Losses in Distribution Networks", **Journal of Water Resources Planning and Management**, 123(5):284-291.

Andersen, J.H. and R.S.Powell (2000) "Implicit state-estimation technique for water network monitoring", **Urban Water** 2, 123-130.

Buchberger, S.G. and L. Wu (1995) "Model for instantaneous residential water demands", **Journal of Hydraulic Engineering**, 121(3):232-246.

Buchberger, S.G., J.T. Carter, Y. Lee and T.G. Schade (2003**) "Random Demands, Travel Times and Water Quality in Deadends of a Municipal Distribution System**", AwwaRF Project Report 294, Denver, Colorado, 476 pages.

Colombo, A.F. and B.W. Karney (2002) "Energy and Costs of Leaky Pipes: Toward Comprehensive Picture", **Journal of Water Resources Planning and Management**, 128(6):441-450.

Hipel, K. W. and A.I. McLeod (1994) **Time Series Modeling of Water Resources and Enviornmental Systems,** Elsevier, Amsterdam, 1013 pages.

Jensen, R. (1987) Newsletters, **Texas Water Resources**, Vol. 13, No.3.

Jowitt, P.W. and C. Xu (1990) "Optimal Valve Control in Water Distribution Networks", **Journal of Water Resources Planning and Management**, 116(4):455-472.

Ligget, J.A. and L-C Chen (1994) "Inverse Transient Analysis in Pipe Networks", **Journal of Hydraulic Engineering**, 120(8):934-955.

Liou, C.P. (1998) "Pipeline Leak Detection by Impulse Response Extraction", **Journal of Fluids Engineering**, Vol. 120. 833-838.

Liou, J.C.P. and J. Tian (1995) "Leak Detection–Transient Flow Simulation Approaches", **Journal of Energy Resources Technology**, Vol. 117, 243-248.

Maslia, M.L., J.B. Sautner, and M.M. Aral (2000) **Analysis of the 1998 Water-Distribution System Serving the Dover Township Area, New Jersey: Field Data Collection Activities and Water-Distribution System Modeling**, Appendix D and E in Supplemental Data, ATSDR, Atlanta, GA.

Mpesha, W., S.L.Gassman and M.H. Chaudhry (2001), "Leak Detection in Pipes by Frequency Response Method", Journal of Hydraulic Engineering, **Journal of Hydraulic Engineering**, 127(2):134-147.

Mpesha, W., S.L.Gassman and M.H. Chaudhry (2002), “Leak Detection in Pipes by Frequency Response Method using a Step Excitation”, **Journal of Hydraulic Research**, Vol. 40. 55-61.

Mukherjee, J. and S. Narasimhan (1996) “Leak Detection in Networks of Pipelines by Generalized Likelihood Ratio Method”, **Industrial and Engineering Chemistry Journa**l, (35) 1886-1893.

Pudar, R.S. and J.A.Ligget (1992) “Leaks in Pipe Networks”, **Journal of Hydraulic Engineering**, 118(7):1031-1046.

Smith, L.A., K.A. Fields, A.S. C. Chen and Anthony N.Tafuri (2000) “**Options for Leak and Break Detection and Repair of Drinking Water Systems**”, Battelle Press, Columbus, 163 pgs.

Stedinger, J.R., R.M. Vogel and E.Foufoula-Georgiou (1993) "Frequency Analyis of Extreme Events" Chapter 18 in **Handbook of Hydrology** edited by D.R. Maidment, McGraw-Hill, New York.

Tucciarelli, T., A.Criminisi and D.Termini (1999) “Leak analysis in pipeline systems by means of optimal valve regulation”, **Journal of Hydraulic Engineering**, 125(3):277-285.

USEPA Fact Sheet, “**1999 Drinking Water Infrastructure Needs Survey”**, U.S. Environmental Protection Agency, EPA-816-01-001, 2001.

Vitkovsky, J.P., A.R. Simpson and M.F. Lambert (1999) "Leak Detection and Calibration of Water Distribution of Systems using Transients and Genetic Algorithms", **Water Distribution Systems Conference, Division of Water Resources Planning and Management**, ASCE, Tempe, Arizona, June 7-9. p 1-10.

Wang, G., D. Dong and C. Fang (1993) "Leak Detection for Transport Pipelines Based on Autoregressive Modeling", **IEEE Transactions on Instrumentation and Measurement,** 42(1):68-71.

http://WWW3.akwien.at/pdf/uv/university_of_greenwich.pdf

http://irc.nrc-cnrc.gc.ca

APPENDIX A

DERIVATION OF MOMENTS FOR MIXED TRUNCATED NORMAL DISTRIBUTION

Moments of a mixed truncated normal distribution are given by

$$E[Q^k] = Q_L^k.\Phi\left(\frac{Q_L - \mu}{\sigma}\right) + \int_{Q_L}^{\infty} \frac{q^k}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right] dq \qquad \text{(A-1)}$$

for k =1 and k = 2, Eq (A-1) gives

$$E[Q] = Q_L.\Phi\left(\frac{Q_L - \mu}{\sigma}\right) + \int_{Q_L}^{\infty} \frac{q}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right] dq \qquad \text{(A-2)}$$

$$E[Q^2] = Q_L^2.\Phi\left(\frac{Q_L - \mu}{\sigma}\right) + \int_{Q_L}^{\infty} \frac{q^2}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right] dq \qquad \text{(A-3)}$$

A.1 MOMENTS OF A MIXED TRUNCATED NORMAL DISTRIBUTION

Evaluation of the definite integral in Eq (A-2)

$$\int_{Q_L}^{\infty} \frac{q}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right] dq$$

$$= \int_{Q_L}^{\infty} \frac{q}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{q^2}{2\sigma^2} + \frac{\mu q}{\sigma} - \frac{\mu^2}{2\sigma^2}\right] dq$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\mu^2}{2\sigma^2}\right] \int_{Q_L}^{\infty} q.\exp\left[-\frac{q^2}{2\sigma^2}\right].\exp\left[\frac{\mu q}{\sigma}\right] dq$$

Let $u = \exp\left[\frac{\mu q}{\sigma^2}\right]$ $\qquad du = \left(\frac{\mu}{\sigma^2}\right)\exp\left[\frac{\mu q}{\sigma^2}\right]dq$

$$dv = q\exp\left[-\frac{q^2}{2\sigma^2}\right]dv \qquad v = (-\sigma^2)\exp\left[-\frac{q^2}{2\sigma^2}\right]$$

Integrating by parts, $\int u dv = uv - \int v du$

Hence, definite integral

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{\exp\left[\frac{\mu q}{\sigma^2}\right](-\sigma^2)\exp\left[-\frac{q^2}{2\sigma^2}\right]_{Q_L}^{\infty} - \int_{Q_L}^{\infty}(-\sigma^2)\exp\left[-\frac{q^2}{2\sigma^2}\right]\left(\frac{\mu}{\sigma^2}\right)\exp\left[\frac{\mu q}{\sigma^2}\right]dq\right\}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{\sigma^2\exp\left[-\frac{Q_L^2}{2\sigma^2} + \frac{2\mu Q_L}{2\sigma^2}\right] + \int_{Q_L}^{\infty}(\mu)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]dq\right\}$$

$$= \sqrt{\frac{\sigma^2}{2\pi}}\exp\left[-\frac{Q_L^2}{2\sigma^2} + \frac{\mu Q_L}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right] + \frac{\mu}{\sqrt{2\pi\sigma^2}}\int_{Q_L}^{\infty}\exp\left[-\frac{q^2}{2\sigma^2} + \frac{\mu q}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right]dq$$

$$= \sqrt{\frac{\sigma^2}{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{Q_L - \mu}{\sigma}\right)^2\right] + \frac{\mu}{\sqrt{2\pi\sigma}}\int_{Q_L}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right]dq$$

$$= \sigma\phi\left(\frac{Q_L - \mu}{\sigma}\right) + \mu\left[1 - \Phi\left(\frac{Q_L - \mu}{\sigma}\right)\right]$$

$$= \sigma\left\{\phi\left(\frac{Q_L - \mu}{\sigma}\right) + \left(\frac{\mu}{\sigma}\right)\left[1 - \Phi\left(\frac{Q_L - \mu}{\sigma}\right)\right]\right\}$$

Now, including the mass spike at Q = Q_L

$$E[Q] = Q_L\Phi\left(\frac{Q_L - \mu}{\sigma}\right) + \sigma\left\{\phi\left(\frac{Q_L - \mu}{\sigma}\right) + \left(\frac{\mu}{\sigma}\right)\left[1 - \Phi\left(\frac{Q_L - \mu}{\sigma}\right)\right]\right\} \qquad \text{(A-4)}$$

Expressing in terms of standardized truncation level, $\varepsilon = \dfrac{Q_L - \mu}{\sigma}$ (A-5)

$$= Q_L \Phi(\varepsilon) + \sigma \left\{ \phi(\varepsilon) + \left(\frac{Q_L}{\sigma} - \varepsilon \right) \Phi(-\varepsilon) \right\}$$

$$= Q_L [\Phi(\varepsilon) + \Phi(-\varepsilon)] + \sigma \{\phi(\varepsilon) - \varepsilon \Phi(-\varepsilon)\}$$

$$= Q_L + \sigma \{\phi(\varepsilon) - \varepsilon \Phi(-\varepsilon)\}$$

Rescaling flow rate, $Q^* = \dfrac{Q - Q_L}{\sigma}$ (A-6)

$$E[Q^*] = \phi(\varepsilon) - \varepsilon \Phi(-\varepsilon) \quad \text{(A-7)}$$

(Eq. 4.5a in the text).

Evaluation of the definite integral in Eq (A-3):

$$\int_{Q_L}^{\infty} \frac{q^2}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2} \left(\frac{q - \mu}{\sigma} \right)^2 \right] dq$$

$$= \int_{Q_L}^{\infty} \frac{q^2}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{q^2}{2\sigma^2} + \frac{\mu q}{\sigma^2} - \frac{\mu^2}{2\sigma^2} \right] dq$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \left(\exp\left[-\frac{\mu^2}{2\sigma^2} \right] \right) \int_{Q_L}^{\infty} q^2 \exp\left[-\frac{q^2}{2\sigma^2} \right] \exp\left[\frac{\mu q}{\sigma^2} \right] dq$$

Let $u = q \exp\left[\dfrac{\mu q}{\sigma^2} \right]$ $\qquad du = \left(\dfrac{\mu q}{\sigma^2} \right) \exp\left[\dfrac{\mu q}{\sigma^2} \right] dq + \exp\left[\dfrac{\mu q}{\sigma^2} \right] dq = \exp\left[\dfrac{\mu q}{\sigma^2} \right] \left(\dfrac{\mu q}{\sigma^2} + 1 \right) dq$

$$dv = q \exp\left[-\frac{q^2}{2\sigma^2} \right] dq \qquad v = (-\sigma^2) \exp\left[-\frac{q^2}{2\sigma^2} \right]$$

Integrating by parts, $\int u dv = uv - \int v du$

Hence, definite integral

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{\begin{array}{l} q\exp\left[\frac{\mu q}{\sigma^2}\right]\left(-\sigma^2\right)\exp\left[-\frac{q^2}{2\sigma^2}\right]_{Q_L}^{\infty} \\ -\int_{Q_L}^{\infty}\left(-\sigma^2\right)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]\left(\frac{\mu q}{\sigma^2}+1\right)dq \end{array}\right\}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{Q_L\sigma^2\exp\left[\frac{\mu Q_L}{\sigma^2}\right]\exp\left[-\frac{Q_L^2}{2\sigma^2}\right]\right\}$$
$$+\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{\int_{Q_L}^{\infty}(\mu q)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]dq+\int_{Q_L}^{\infty}\left(\sigma^2\right)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]dq\right\}$$

$$= \left(\frac{Q_L\sigma}{\sqrt{2\pi}}\right)\left\{\exp\left[-\frac{Q_L^2}{2\sigma^2}+\frac{\mu Q_L}{\sigma^2}-\frac{\mu^2}{2\sigma^2}\right]\right\}$$
$$+\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{\int_{Q_L}^{\infty}(\mu q)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]dq+\int_{Q_L}^{\infty}\left(\sigma^2\right)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]dq\right\}$$

$$= Q_L\sigma\phi\left(\frac{Q_L-\mu}{\sigma}\right)$$
$$+\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{\int_{Q_L}^{\infty}(\mu q)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]dq\right\}$$
$$+\left(\frac{\sigma^2}{\sqrt{2\pi\sigma^2}}\right)\left\{\int_{Q_L}^{\infty}\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right\}dq$$

$$= Q_L\sigma\phi\left(\frac{Q_L-\mu}{\sigma}\right)$$
$$+\left(\frac{\mu}{\sqrt{2\pi\sigma^2}}\right)\left(\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\right)\left\{\int_{Q_L}^{\infty}(q)\exp\left[-\frac{q^2}{2\sigma^2}\right]\exp\left[\frac{\mu q}{\sigma^2}\right]dq\right\}$$
$$+\left(\frac{\sigma^2}{\sqrt{2\pi\sigma^2}}\right)\left\{\int_{Q_L}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{q-\mu}{\sigma}\right)^2\right]\right\}dq$$

Evaluating the second term from previous integral and including mass spike at Q = Q_L

$$= Q_L^2 \Phi\left(\frac{Q_L - \mu}{\sigma}\right) + Q_L \sigma \phi\left(\frac{Q_L - \mu}{\sigma}\right) + \mu\sigma\left\{\phi\left(\frac{Q_L - \mu}{\sigma}\right) + \left(\frac{\mu}{\sigma}\right)\left[1 - \Phi\left(\frac{Q_L - \mu}{\sigma}\right)\right]\right\} + \sigma^2\left[1 - \Phi\left(\frac{Q_L - \mu}{\sigma}\right)\right]$$

$$E\left[Q^2\right] = Q_L^2 \Phi\left(\frac{Q_L - \mu}{\sigma}\right) + \sigma^2\left\{\left(\frac{Q_L + \mu}{\sigma}\right)\phi\left(\frac{Q_L - \mu}{\sigma}\right) + \left(1 + \frac{\mu^2}{\sigma^2}\right)\left[1 - \Phi\left(\frac{Q_L - \mu}{\sigma}\right)\right]\right\} \qquad \text{(A-8)}$$

Again, expressing in terms of standardized truncation level, $\varepsilon = \dfrac{Q_L - \mu}{\sigma}$

$$= Q_L^2 \Phi(\varepsilon) + (Q_L + \mu)\sigma\phi(\varepsilon) + \left(\mu^2 + \sigma^2\right)\Phi(-\varepsilon)$$

$$= \sigma^2\left\{\left(\frac{Q_L^2}{\sigma^2}\right)\Phi(\varepsilon) + \left(\frac{Q_L}{\sigma} + \frac{\mu}{\sigma}\right)\phi(\varepsilon) + \left(\frac{\mu^2}{\sigma^2} + 1\right)\Phi(-\varepsilon)\right\}$$

$$= \sigma^2\left\{\left(\frac{Q_L^2}{\sigma^2}\right)\Phi(\varepsilon) + \left(\frac{2Q_L}{\sigma} - \varepsilon\right)\phi(\varepsilon) + \left(\frac{Q_L^2}{\sigma^2} - \frac{2Q_L\varepsilon}{\sigma} + \varepsilon^2 + 1\right)\Phi(-\varepsilon)\right\}$$

$$= Q_L^2 + \sigma^2\left\{\left(\frac{2Q_L}{\sigma} - \varepsilon\right)\phi(\varepsilon) + \left(-\frac{2Q_L\varepsilon}{\sigma} + \varepsilon^2 + 1\right)\Phi(-\varepsilon)\right\}$$

Rescaling flow rates, $Q^* = \dfrac{Q - Q_L}{\sigma}$

$$E\left[\left(Q^*\right)^2\right] = -\varepsilon\phi(\varepsilon) + \left(\varepsilon^2 + 1\right)\Phi(-\varepsilon) \qquad \text{(A-9)}$$

Noting that $Var[Q] = E[Q^2] - E[Q]^2$, variance of rescaled flow rates can be obtained from Eqs. (7) and (9).

$$Var[Q^*] = [1 + 2\varepsilon\phi(\varepsilon) + \varepsilon^2\Phi(\varepsilon)]\Phi(-\varepsilon) - [\varepsilon + \phi(\varepsilon)]\phi(\varepsilon) \quad \text{(A-10)}$$

(Eq. 4.5b in the text).

A.2 DERIVATIVES OF MOMENTS

Derivatives of EQ*] and S[Q*] are obtained from Eqs. (7) and (10).

$$\frac{dE[Q^*]}{d\varepsilon} = E'[Q^*] = \frac{d}{d\varepsilon}[\phi(\varepsilon) - \varepsilon\Phi(-\varepsilon)]$$

Noting that $\phi'(\varepsilon) = -\varepsilon\phi(\varepsilon), \Phi'(\varepsilon) = \phi(\varepsilon)$ and $\Phi(-\varepsilon) = 1 - \Phi(\varepsilon)$

$$E'[Q^*] = \phi'(\varepsilon) - \frac{d}{d\varepsilon}[\varepsilon - \varepsilon\Phi(\varepsilon)]$$

$$= -\varepsilon\phi(\varepsilon) - [1 - \varepsilon\phi(\varepsilon) - \Phi(\varepsilon)]$$

$$= -[1 - \Phi(\varepsilon)]$$

$$E'[Q^*] = -\Phi(-\varepsilon) \quad \text{(A-11)}$$

(Eq. 4.6a in the text).

$$S'[Q^*] = \frac{d}{d\varepsilon}\left(Var[Q^*]\right)^{1/2} = \frac{\frac{d}{d\varepsilon}\left(Var[Q^*]\right)}{2\sqrt{Var[Q^*]}}$$

Evaluation of numerator

$$\frac{d}{d\varepsilon}\left(Var\left[Q^*\right]\right)=\frac{d}{d\varepsilon}\left(E\left[\left(Q^*\right)^2\right]\right)-\frac{d}{d\varepsilon}\left(E\left[Q^*\right]\right)^2$$

$$=\frac{d}{d\varepsilon}\left\{-\varepsilon\phi(\varepsilon)+\left(1+\varepsilon^2\right)\Phi(-\varepsilon)\right\}-\frac{d}{d\varepsilon}\left\{\phi(\varepsilon)-\varepsilon\Phi(-\varepsilon)\right\}^2$$

$$=\varepsilon^2\phi(\varepsilon)-\phi(\varepsilon)-\phi(\varepsilon)+2\varepsilon\Phi(-\varepsilon)-\varepsilon^2\phi(\varepsilon)-2\left\{\phi(\varepsilon)-\varepsilon\Phi(-\varepsilon)\right\}\left[-\varepsilon\phi(\varepsilon)-\Phi(-\varepsilon)+\varepsilon\phi(\varepsilon)\right]$$

$$=-2\left[\phi(\varepsilon)-\varepsilon\Phi(-\varepsilon)\right]\Phi(\varepsilon)$$

Hence,

$$S'\left[Q^*\right]=\frac{\left[\phi(\varepsilon)-\varepsilon\Phi(-\varepsilon)\right]\Phi(\varepsilon)}{\sqrt{\left[1+2\varepsilon\phi(\varepsilon)+\varepsilon^2\Phi(\varepsilon)\right]\Phi(-\varepsilon)-\left[\varepsilon+\phi(\varepsilon)\right]\phi(\varepsilon)}}$$

$$=\beta(\varepsilon)\Phi(\varepsilon)$$

where, $\beta(\varepsilon)=\dfrac{\left[\phi(\varepsilon)-\varepsilon\Phi(-\varepsilon)\right]}{\sqrt{\left[1+2\varepsilon\phi(\varepsilon)+\varepsilon^2\Phi(\varepsilon)\right]\Phi(-\varepsilon)-\left[\varepsilon+\phi(\varepsilon)\right]\phi(\varepsilon)}}=\dfrac{-1}{\sqrt{\dfrac{E\left[\left(Q^*\right)^2\right]}{E\left[Q^*\right]^2}-1}}$

Therefore,

$$S'\left[Q^*\right]=-\frac{\Phi(\varepsilon)}{\Theta(\varepsilon)} \tag{A-12}$$

(Eq. 4.6b in the text).

where $\Theta(\varepsilon)=\dfrac{S[Q^*]}{E[Q^*]}$ is the coefficient of variation of the rescaled flow rates.

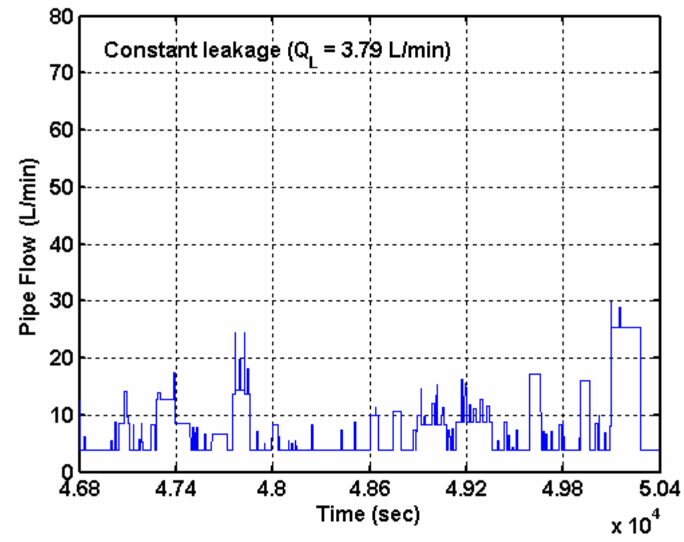
APPENDIX B1

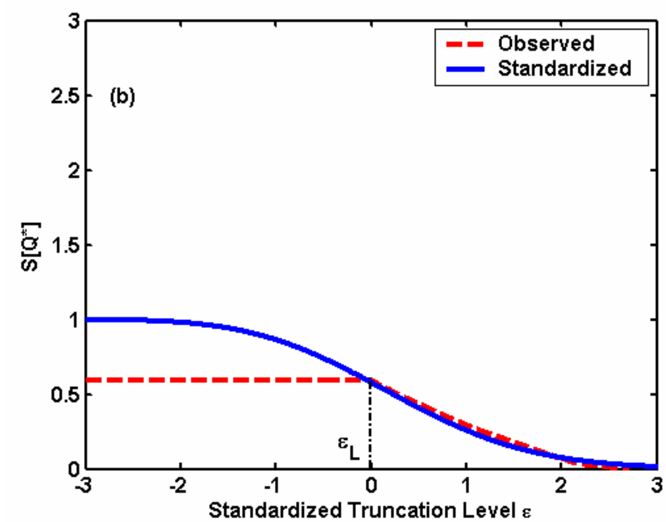
OBSERVED AND STANDARDIZED CURVES FOR CONSTANT LEAKAGE CASE

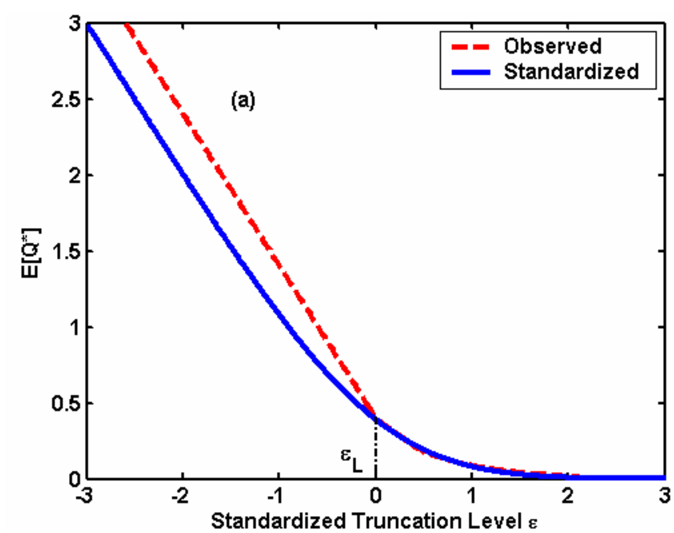
The Figures shown in this Appendix are examples for leak estimation for the constant leakage case listed in Table 5.1 of the text. They include

1. Moments of integrated flows of measured water demands of a distribution site in Milford, Ohio (7 data sets from April through October) and
2. Moments of simulated flows obtained from PRP generator.

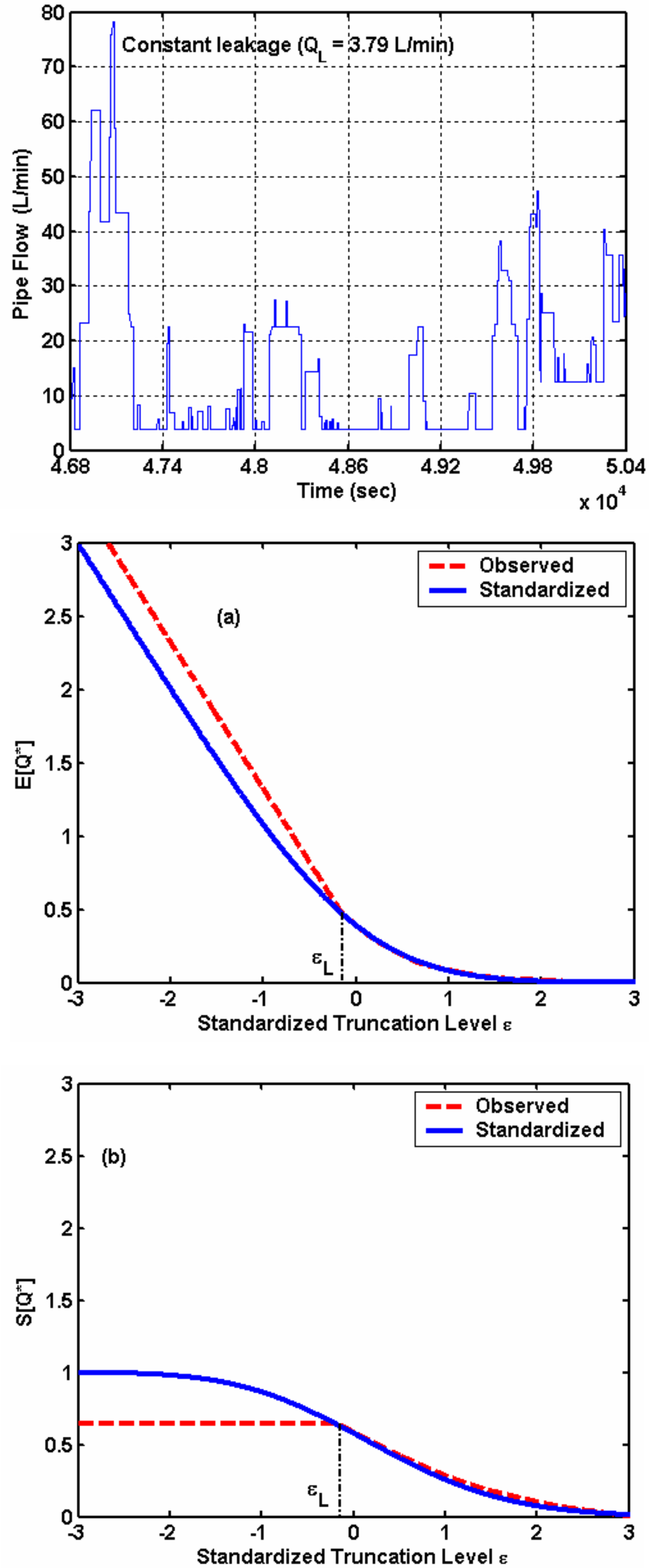
These plots can give a better understanding of the procedure employed to identify the point of departure between the curves of computed sample statistics and those of the standardized moments, and thus estimate leakage rates.



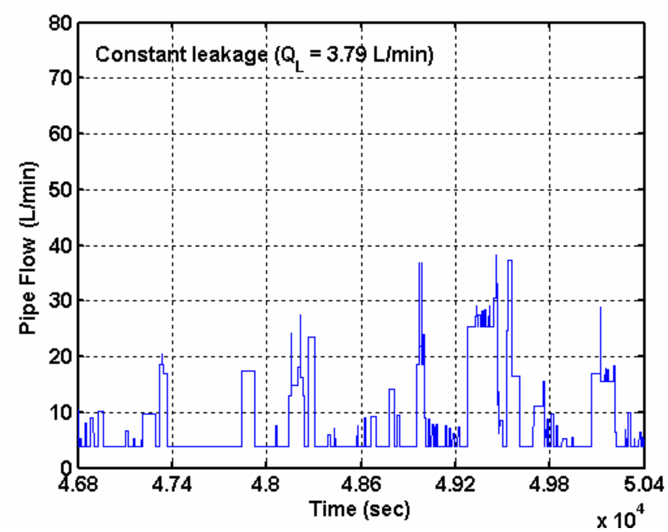


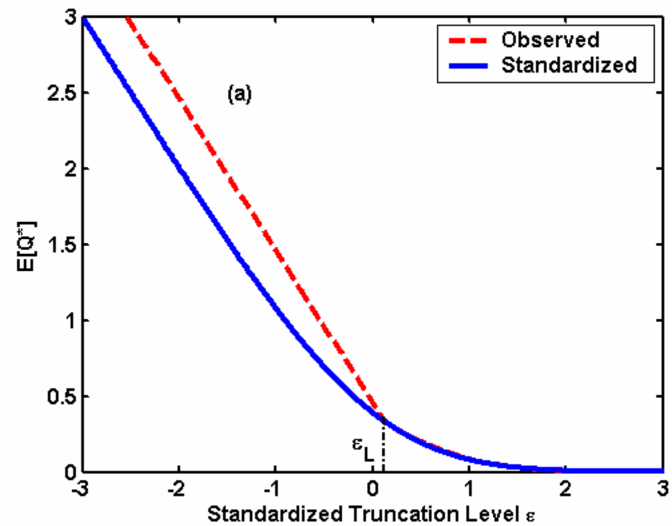


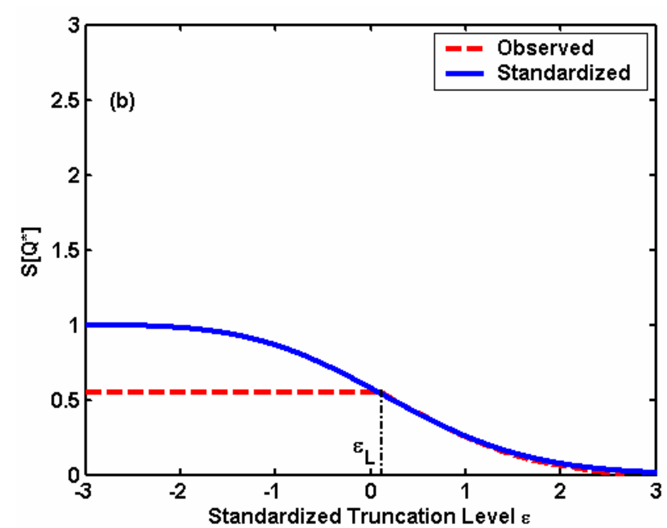
April 01, 1997


Constant leakage (Q_L = 3.79 L/min)
Pipe Flow (L/min)
Time (sec)
x 10^4
4.68
4.74
4.8
4.86
4.92
4.98
5.04
0
10
20
30
40
50
60
70
80
Observed
Standardized
(a)
E[Q*]
ε_L
Standardized Truncation Level ε
(b)
S[Q*]
ε_L
Standardized Truncation Level ε

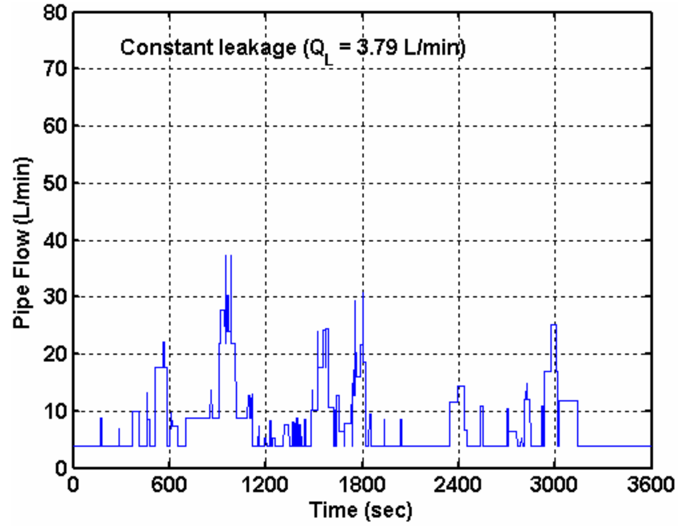
May 01, 1997

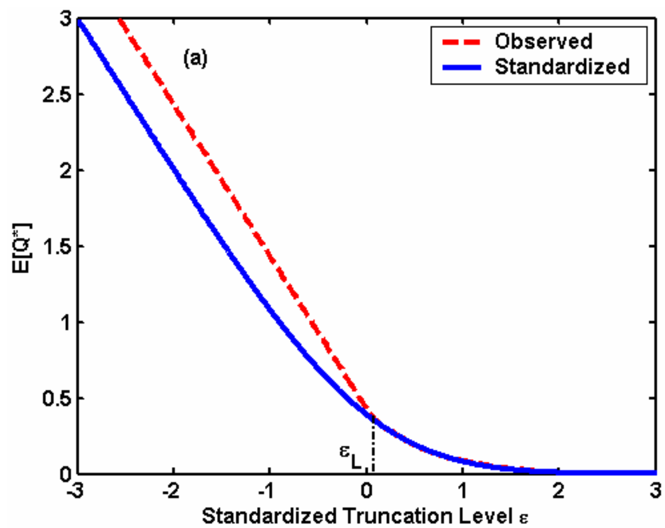

Constant leakage (Q_L = 3.79 L/min)
Pipe Flow (L/min)
Time (sec)
x 10^4

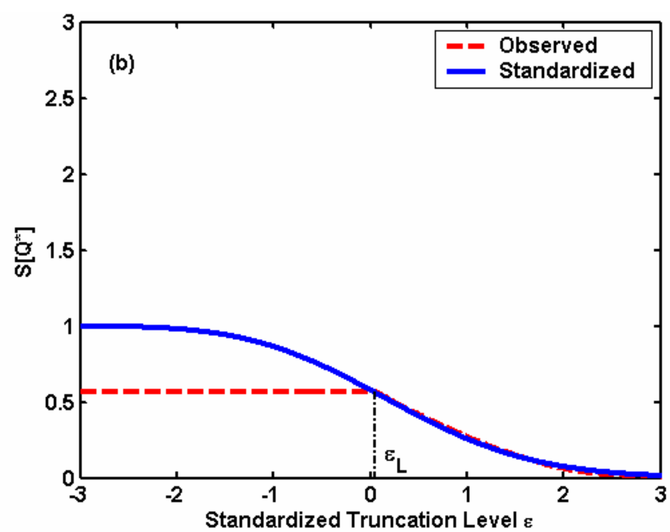

Observed
Standardized
(a)
E[Q*]
ε_L
Standardized Truncation Level ε


Observed
Standardized
(b)
S[Q*]
ε_L
Standardized Truncation Level ε

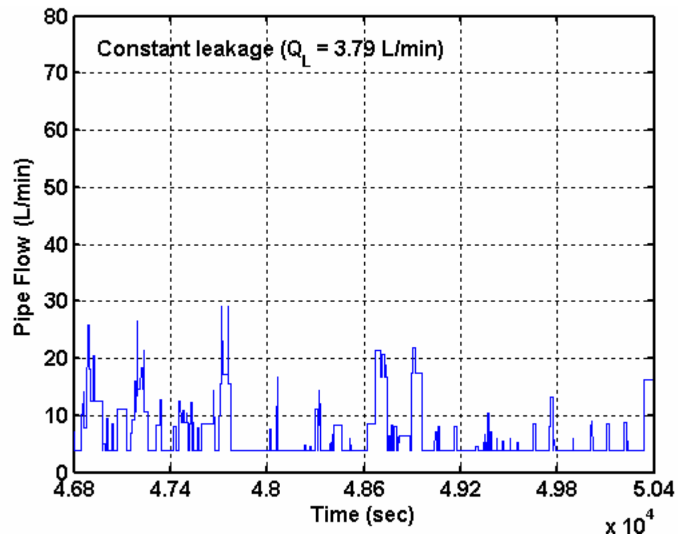
June 03, 1997

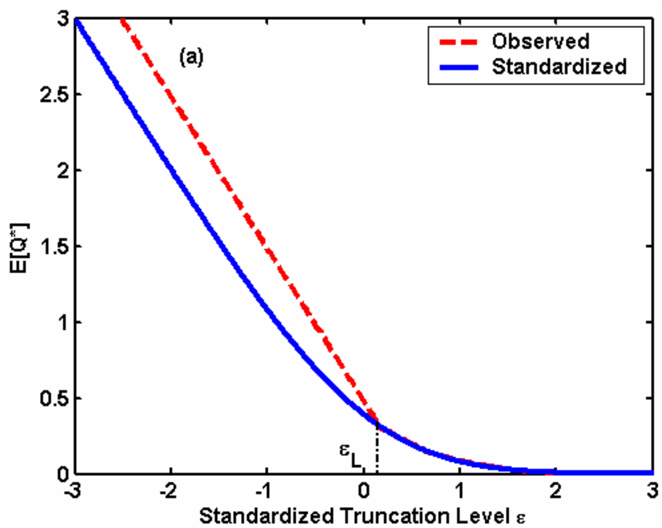

Constant leakage (Q_L = 3.79 L/min)
Pipe Flow (L/min)
Time (sec)

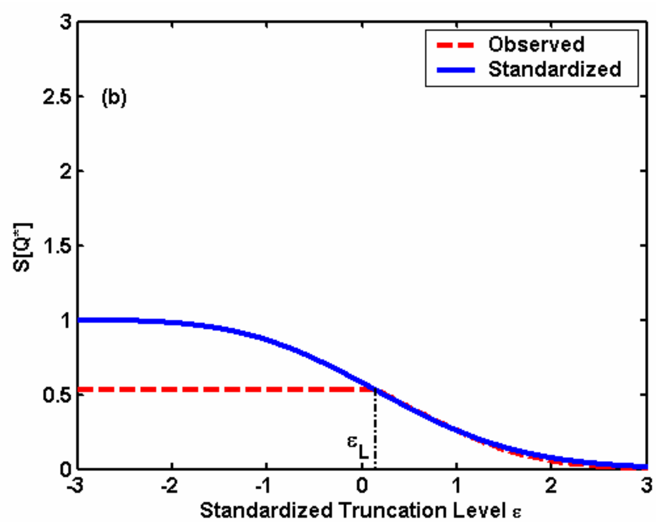

(a)
Observed
Standardized
E[Q*]
ε_L
Standardized Truncation Level ε


(b)
Observed
Standardized
S[Q*]
ε_L
Standardized Truncation Level ε

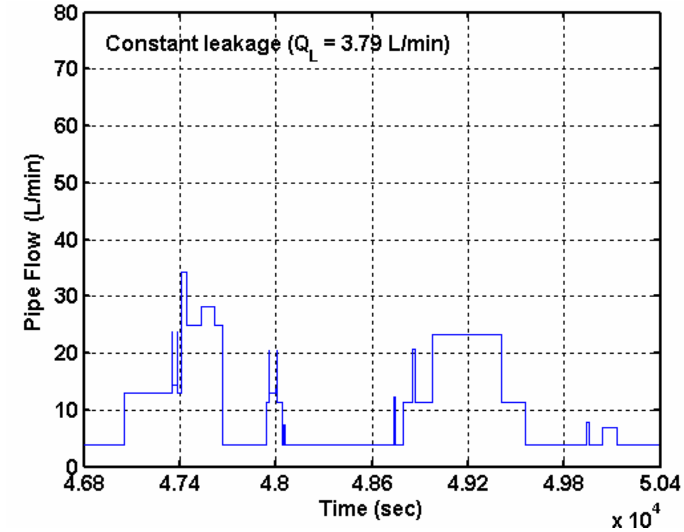
July 25, 1997

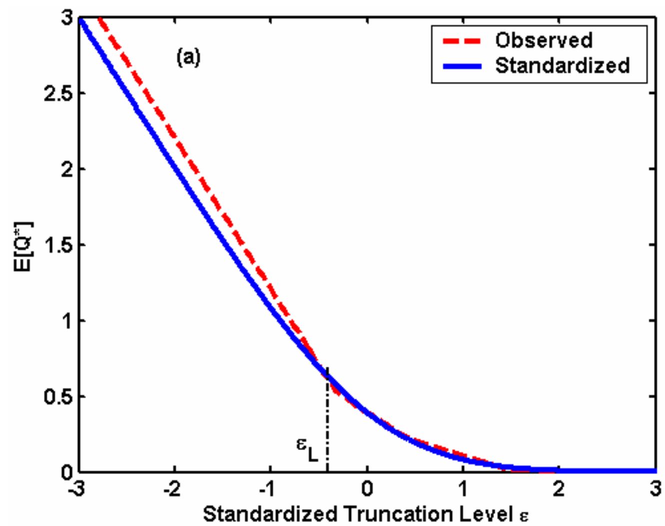


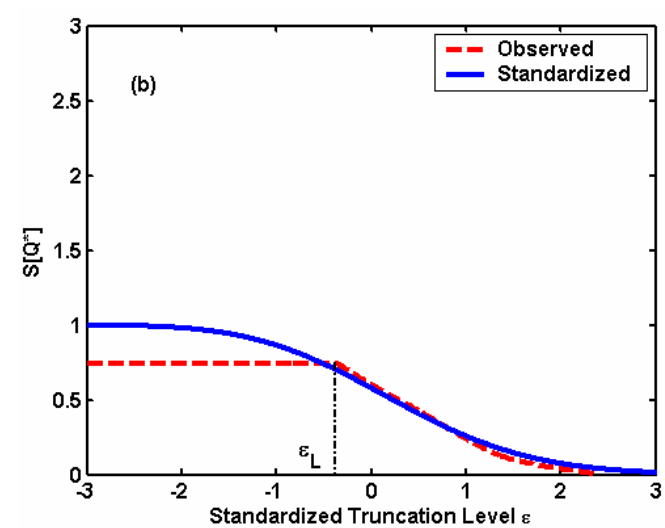




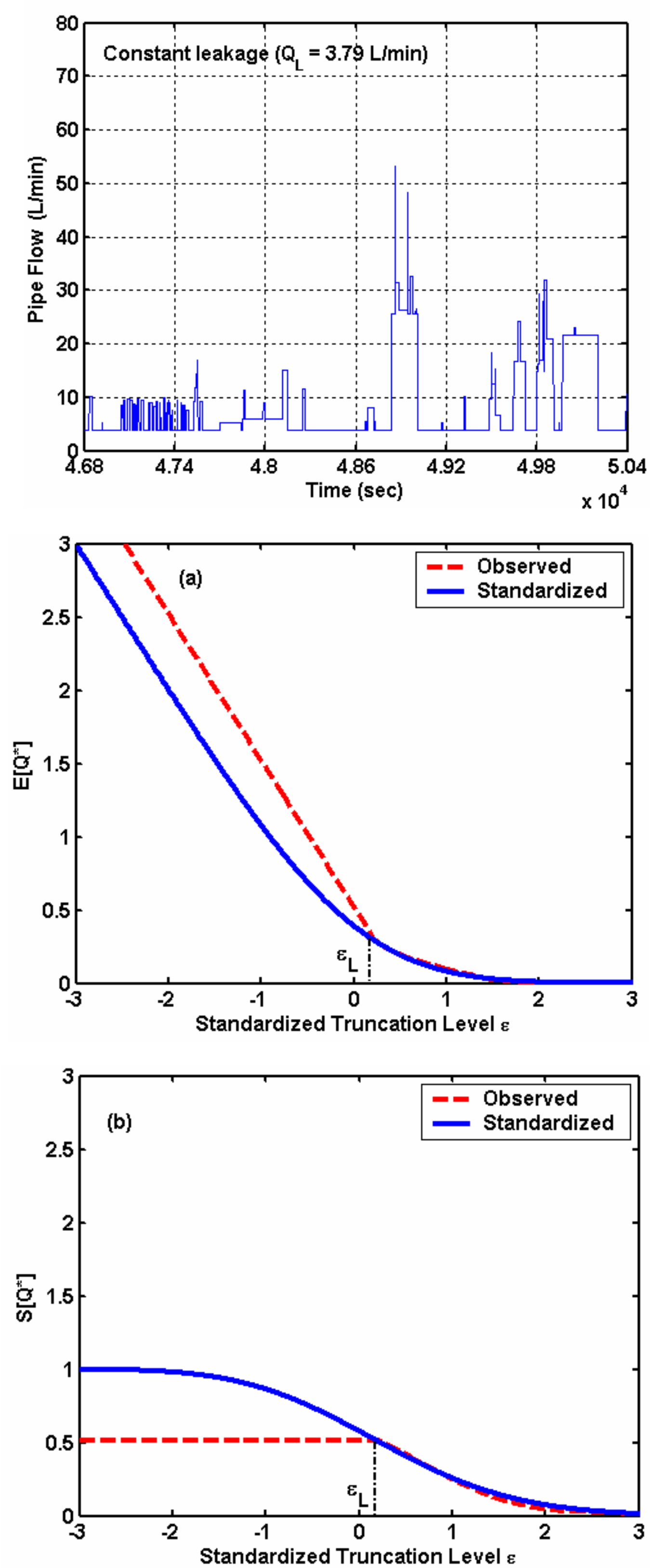
August 05, 1997



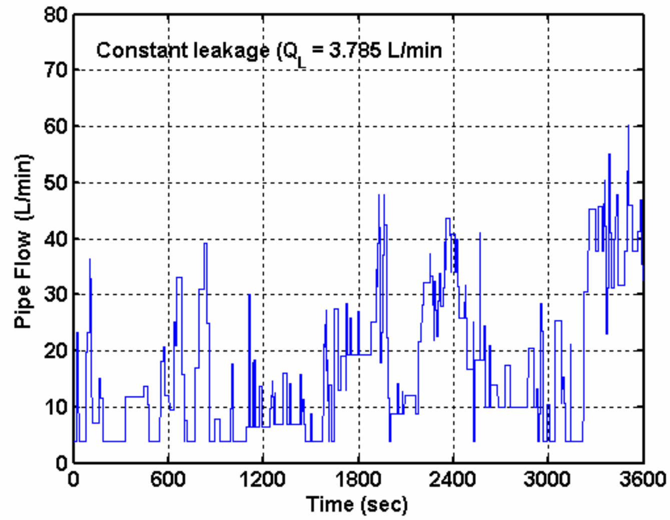


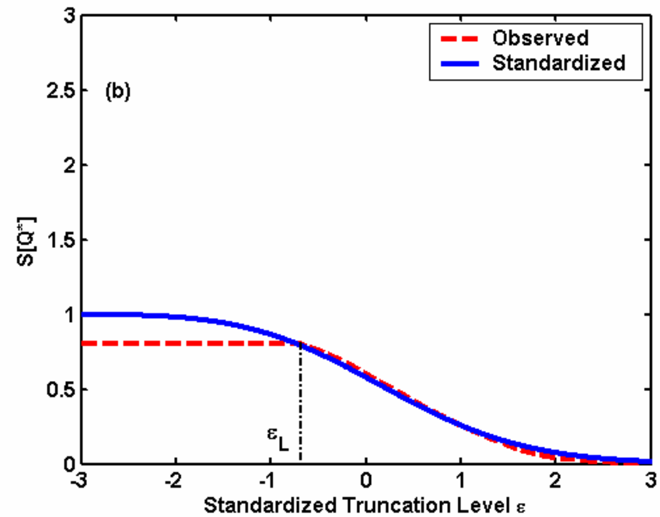


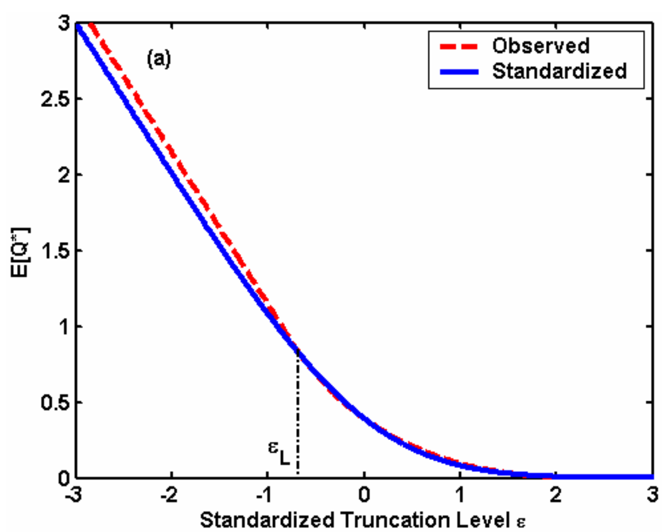
September 01, 1997



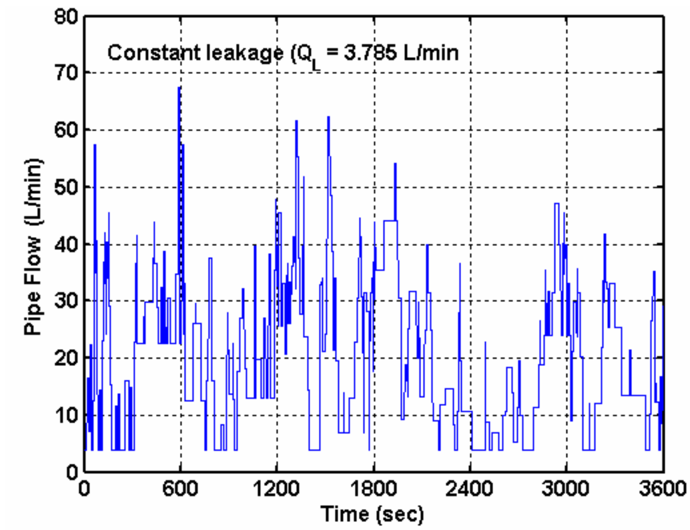
October 16, 1997

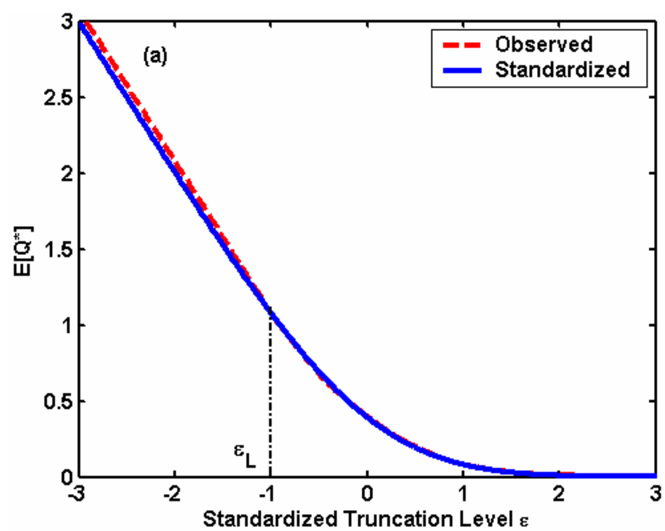


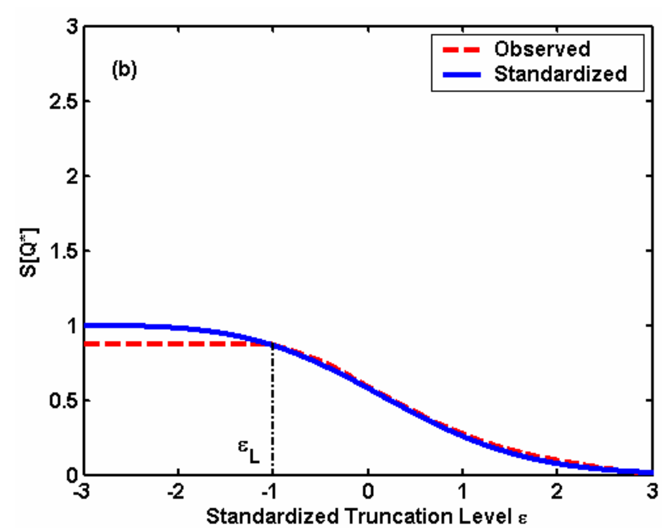




Simulated flows, 400 homes





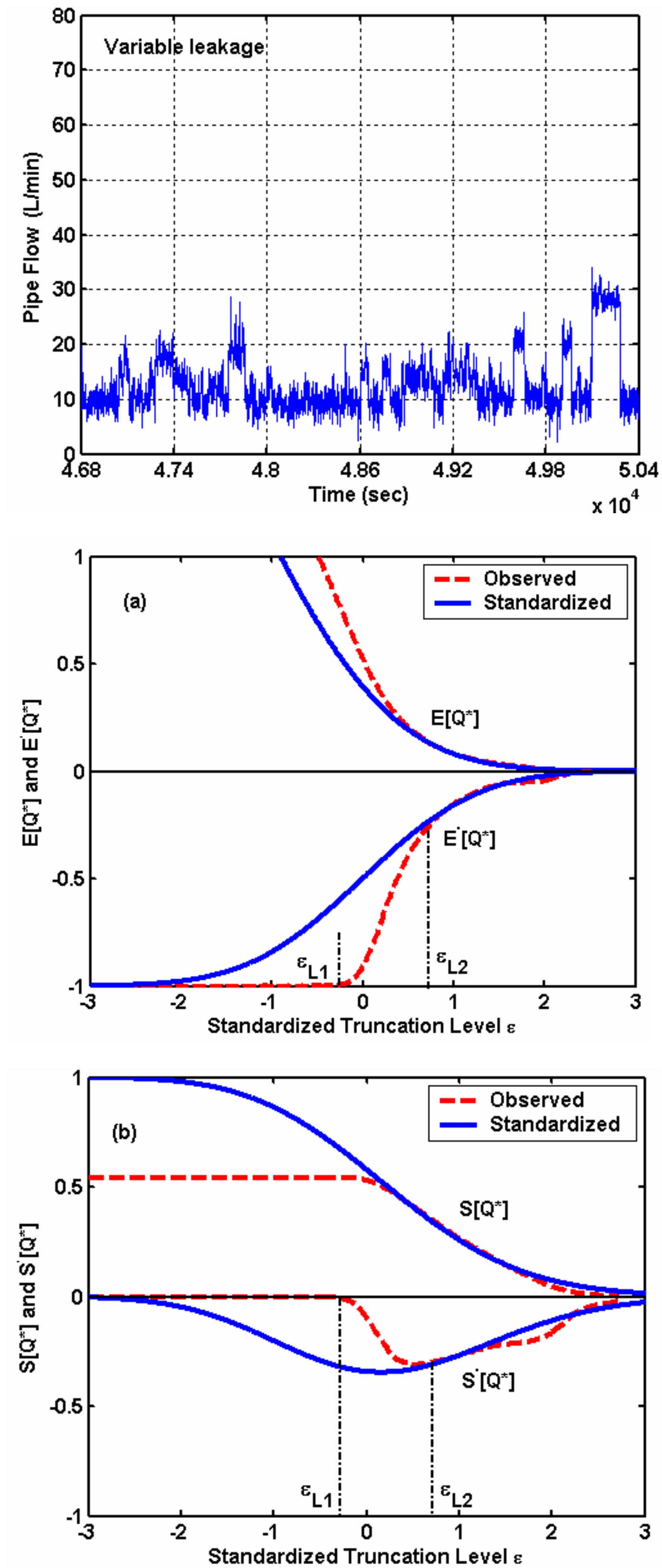


Simulated flows, 500 Homes

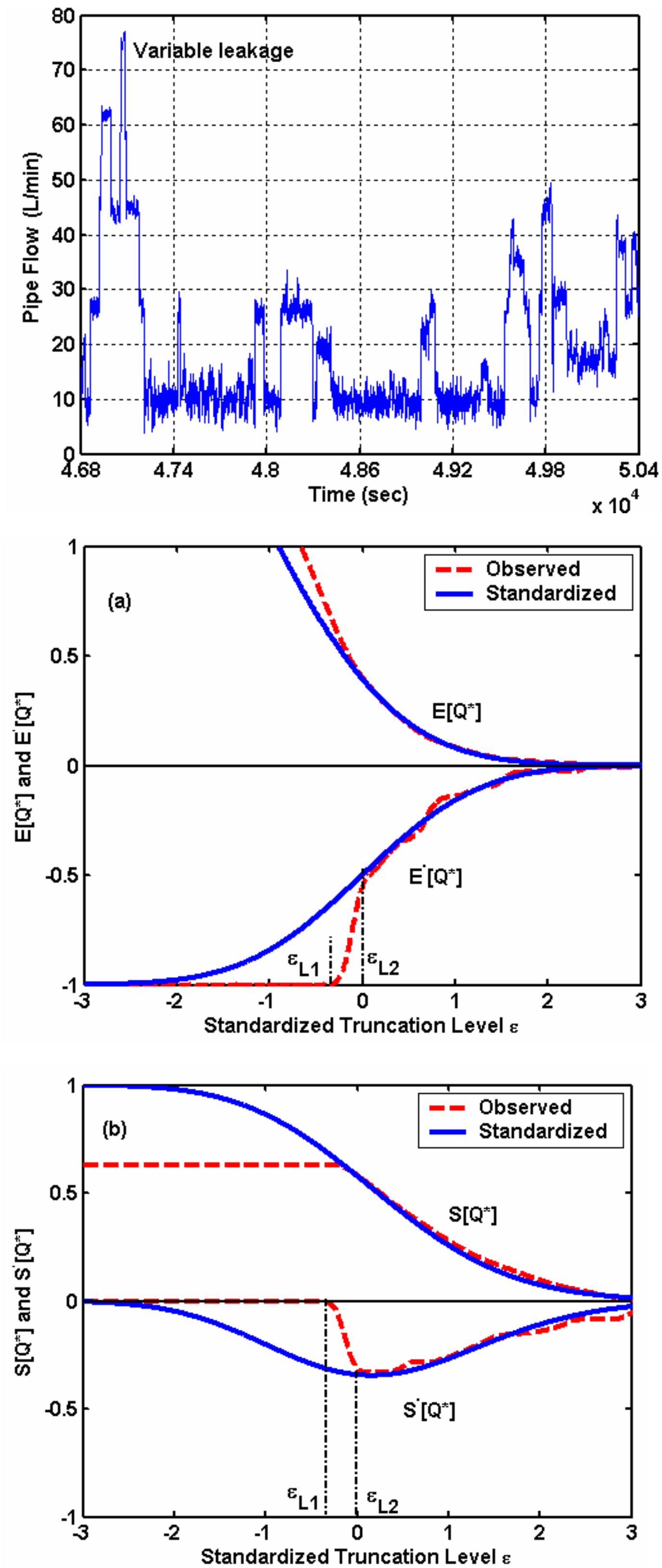
APPENDIX B2

OBSERVED AND STANDARDIZED CURVES FOR VARIABLE LEAKAGE CASE

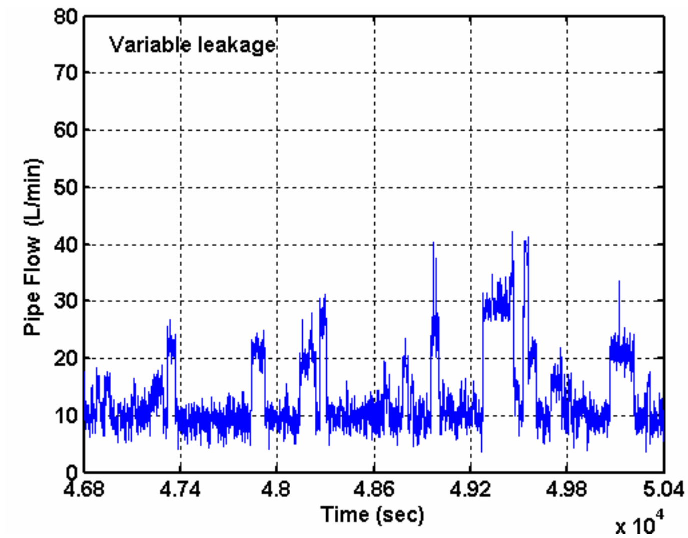
Figures for variable leakage case examples listed in Table 5.2 are included in this part of the Appendix.

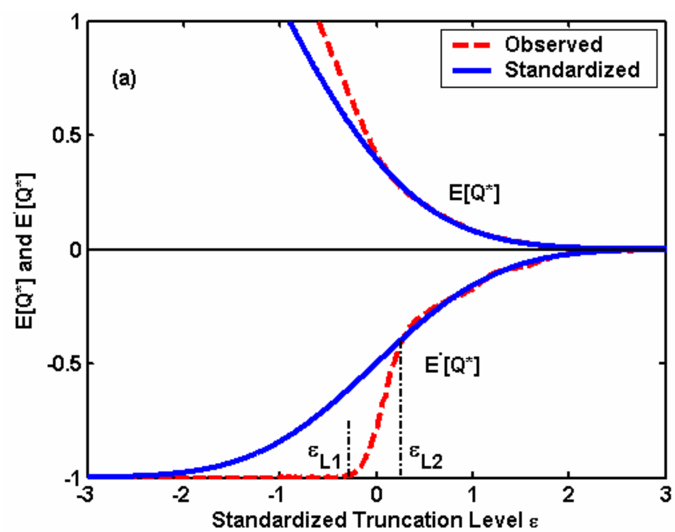


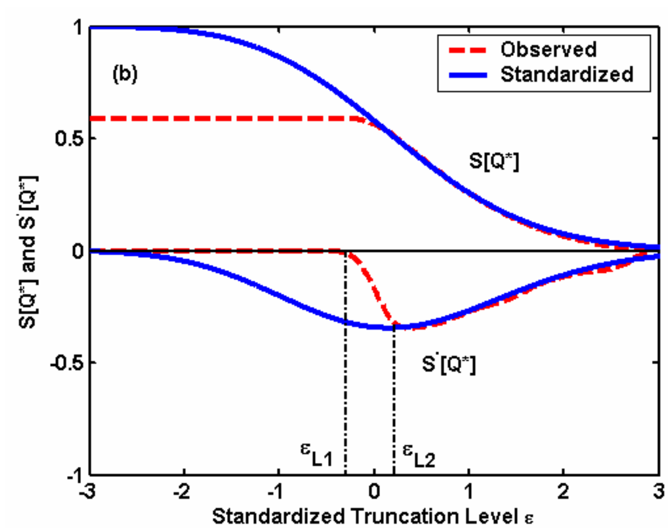
April 01, 1997


Variable leakage
Pipe Flow (L/min)
Time (sec)
x 10^4
(a)
Observed
Standardized
E[Q*]
E-[Q*]
E[Q*] and E-[Q*]
εL1
εL2
Standardized Truncation Level ε
(b)
Observed
Standardized
S[Q*]
S-[Q*]
S[Q*] and S-[Q*]
εL1
εL2
Standardized Truncation Level ε

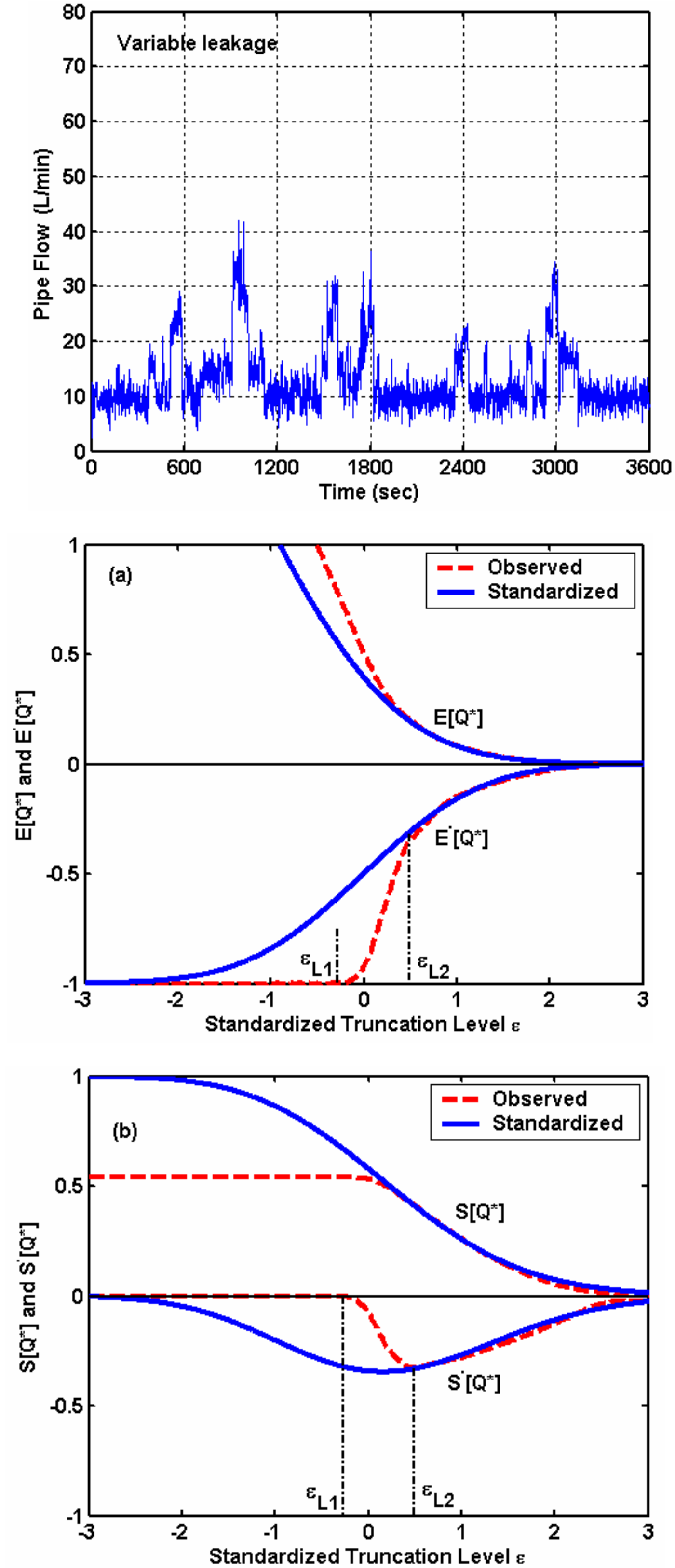
May 01, 1997


Variable leakage
Pipe Flow (L/min)
Time (sec)
x 10^4
80
70
60
50
40
30
20
10
0
4.68
4.74
4.8
4.86
4.92
4.98
5.04

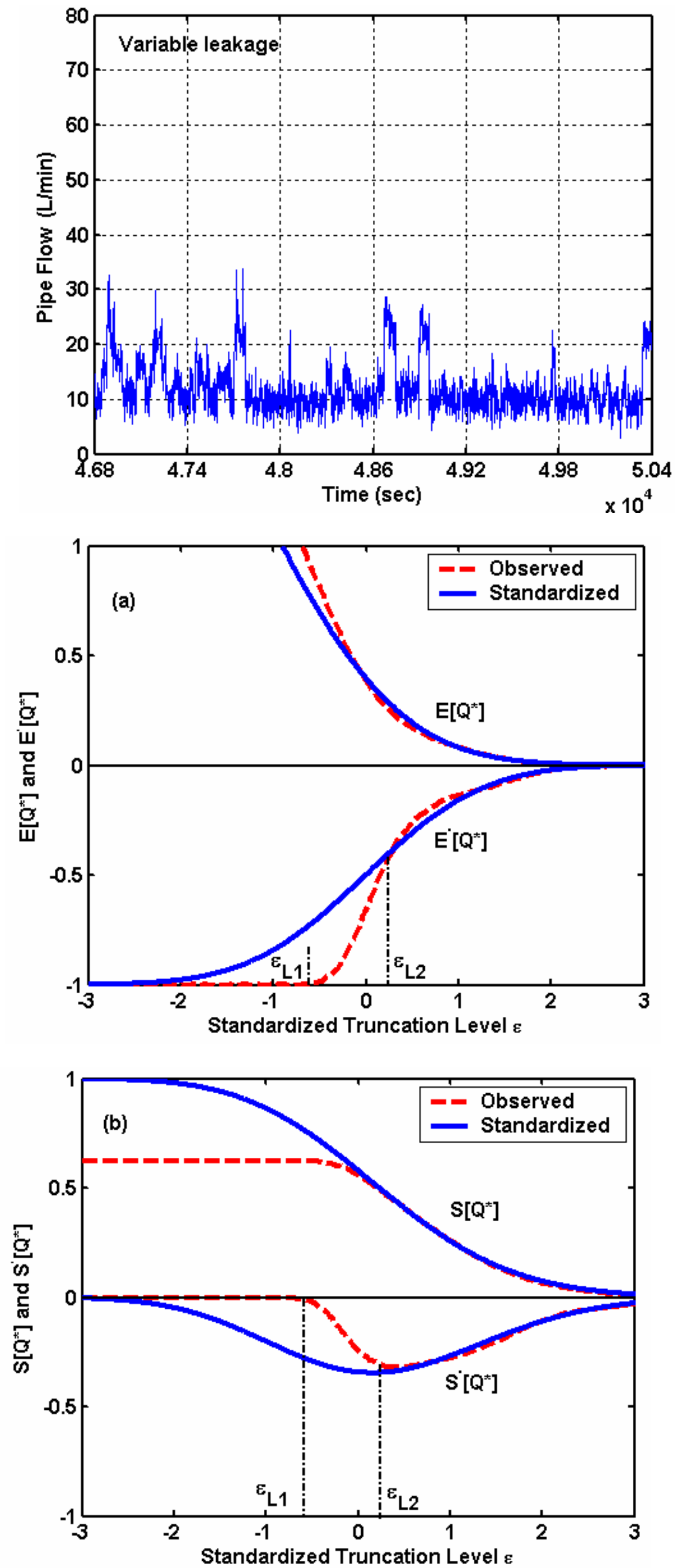

(a)
Observed
Standardized
E[Q*]
E'[Q*]
εL1
εL2
E[Q*] and E'[Q*]
Standardized Truncation Level ε
1
0.5
0
-0.5
-1
-3
-2
-1
0
1
2
3


(b)
Observed
Standardized
S[Q*]
S'[Q*]
εL1
εL2
S[Q*] and S'[Q*]
Standardized Truncation Level ε
1
0.5
0
-0.5
-1
-3
-2
-1
0
1
2
3

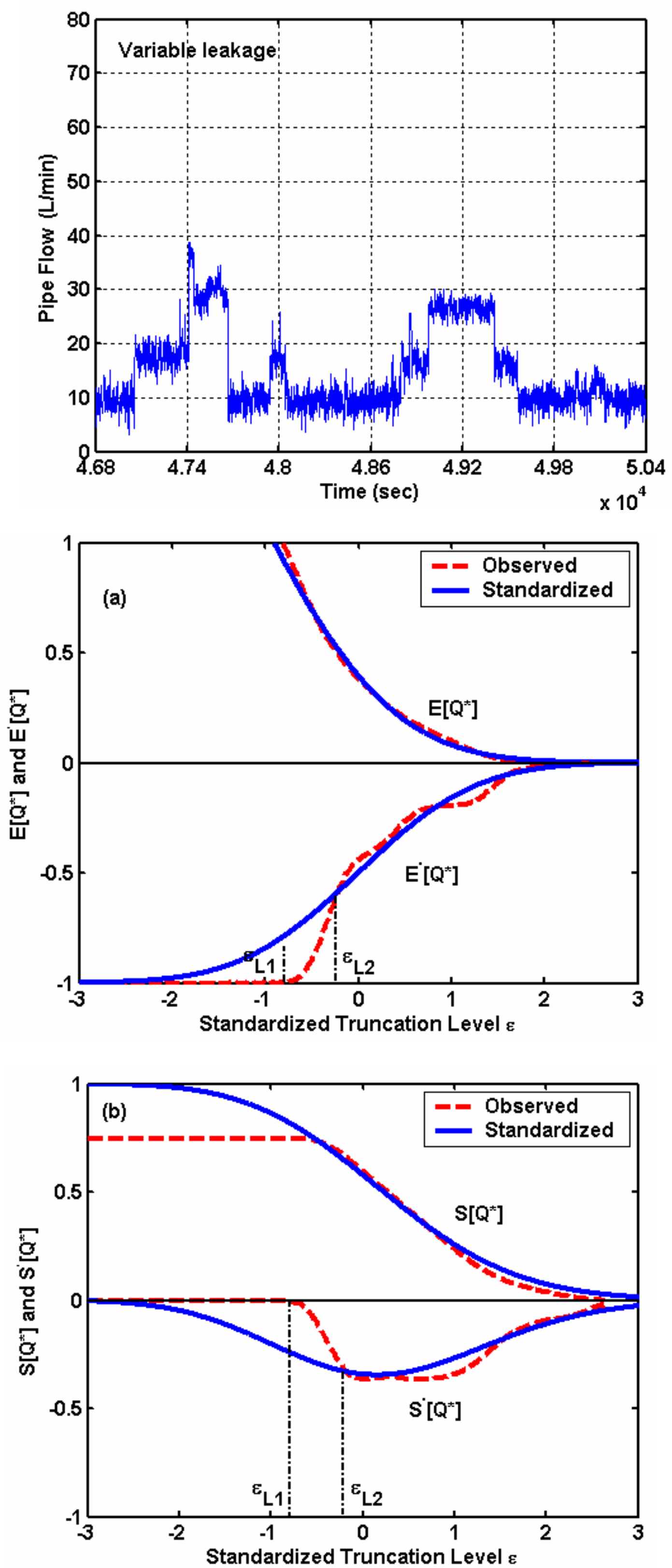
June 02, 1997


Variable leakage
Pipe Flow (L/min)
80
70
60
50
40
30
20
10
0
0
600
1200
1800
2400
3000
3600
Time (sec)
(a)
Observed
Standardized
E[Q*]
E⁻[Q*]
E[Q*] and E⁻[Q*]
1
0.5
0
-0.5
-1
εL1
εL2
-3
-2
-1
0
1
2
3
Standardized Truncation Level ε
(b)
Observed
Standardized
S[Q*]
S⁻[Q*]
S[Q*] and S⁻[Q*]
1
0.5
0
-0.5
-1
εL1
εL2
-3
-2
-1
0
1
2
3
Standardized Truncation Level ε

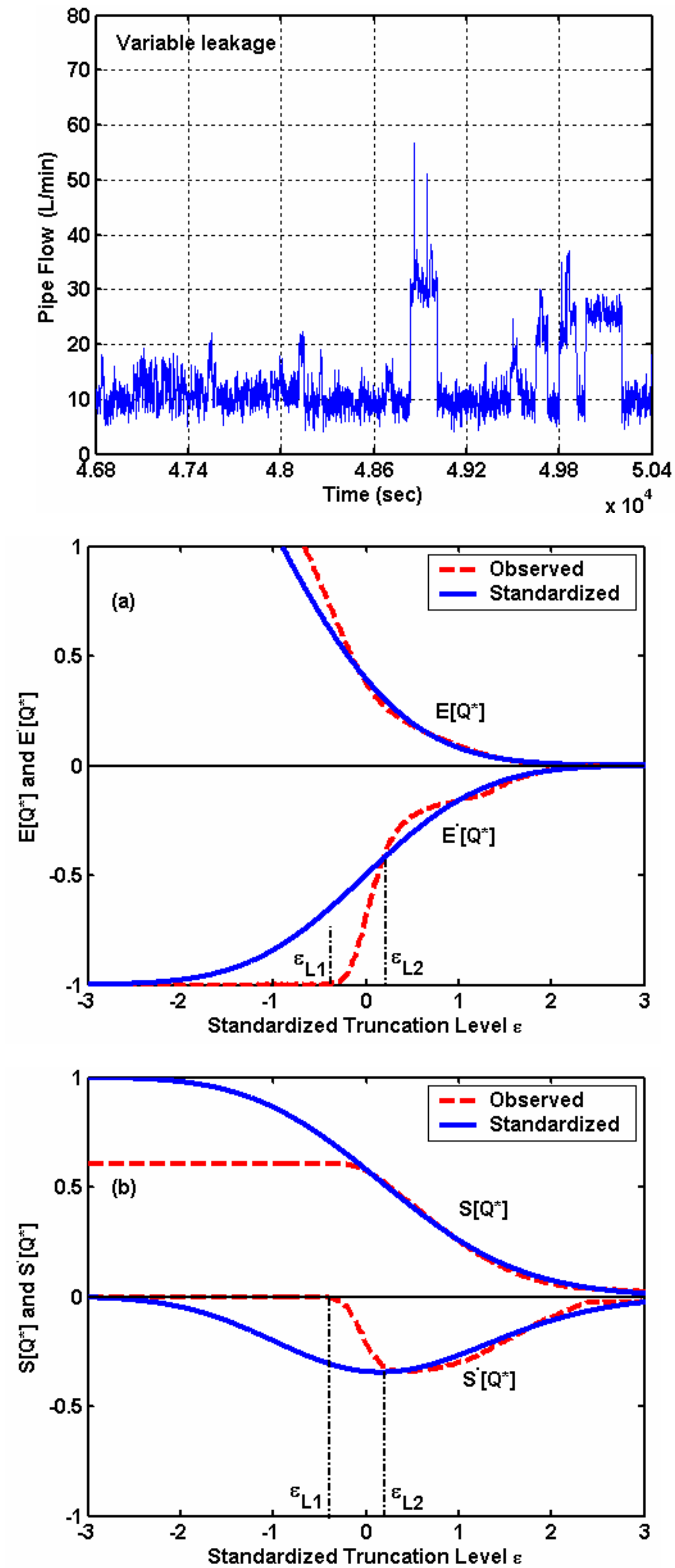
July 25, 1997



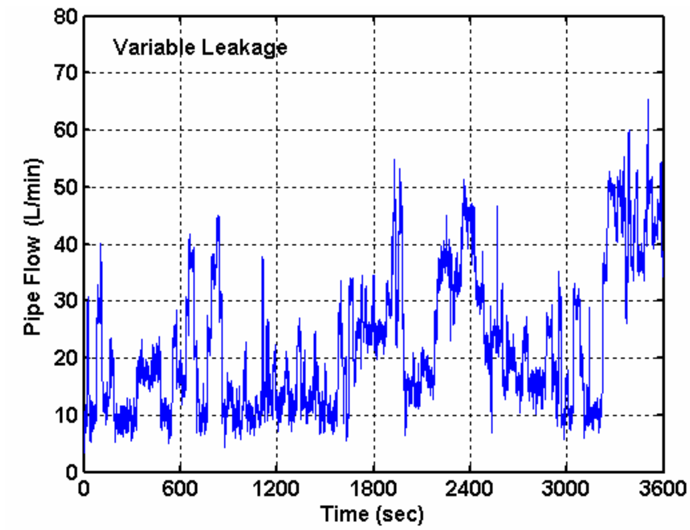
August 05, 1997

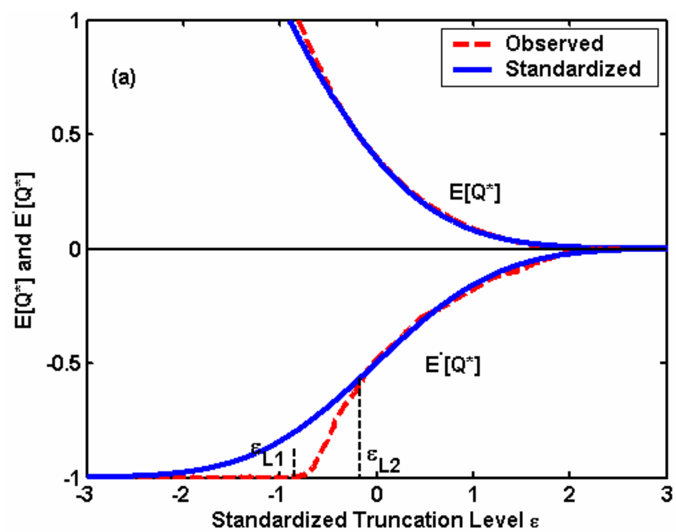


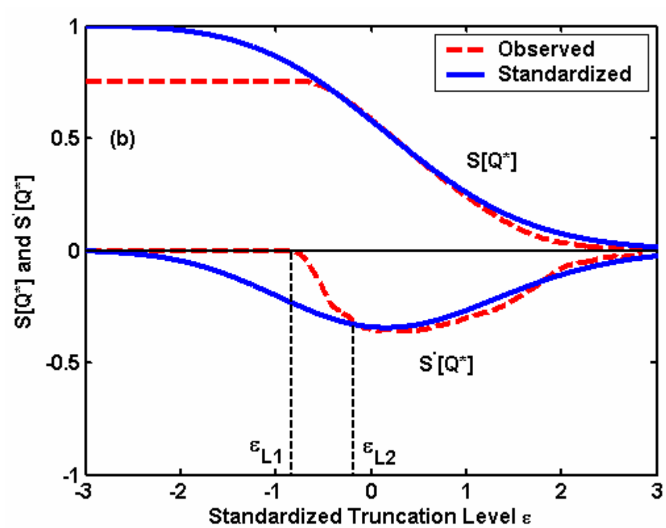
September 01, 1997



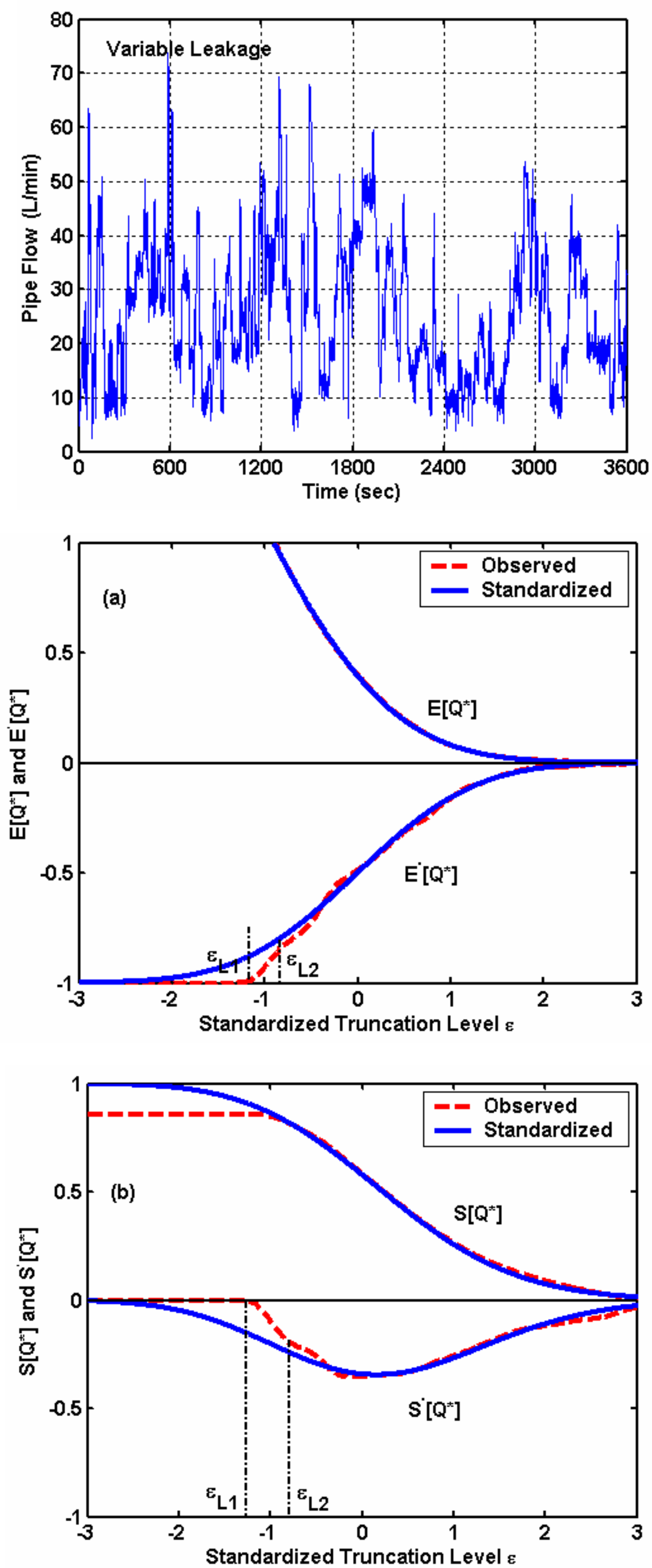
October 16, 1997







Simulated flows, 400 homes



Simulated flows, 500 homes

APPENDIX C

ESTIMATION OF LEAKAGE RATES FOR MEASURED FLOWS IN POLAND

The leak detection algorithm was tested on real flow data obtained from Poland. Pipe flows were recorded continuously for three days on a water main supplying an apartment complex of 2000 flats. Leak analysis was conducted on the recorded pipe flows. An observation period of one hour was chosen from a period of low water usage.

It is quite evident from the time series plot of original data that a background leakage rate of approximately 0.5 m^3/h is present in the flows (see Figure C1). This fact is also confirmed from the vertical mass spike at 0.5 m^3/h of the normal probability plot (see Figure C3). In the absence of pipe leakage, the vertical mass spike would center at zero, suggesting a stagnation frequency of around 10 to 15 %.

The continuous part of the normal probability plot shows a slight curvature, and hence, the data were normalized using a Box-Cox transform. The following two Equations were used for the transformation.

$$x(\lambda) = \frac{\left(x^{\lambda} - 1\right)}{\lambda} \quad \lambda \neq 0$$

$$x(\lambda) = \ln(x) \quad \lambda = 0 \tag{C-1}$$

For a vector of flow data given by $\mathbf{x} = x_1, x_2, \ldots x_n$, the power λ is chosen that maximizes the logarithm of likelihood function

$$f(\mathbf{x},\lambda) = -\frac{n}{2}\ln\left[\sum_{i=1}^{n}\frac{\left(x_i(\lambda)-\bar{x}(\lambda)\right)^2}{n}\right] + (\lambda-1)\sum_{i=1}^{n}\ln x_i$$

where (C-2)

$$\bar{x}(\lambda) = \frac{1}{n}\sum_{i=1}^{n} x_i(\lambda)$$

is the arithmetic mean of the transformed data.

Some of the transformed data were negative. This is not suitable for leak analysis. Hence, absolute minimum negative value was added to each of the transformed flows to ensure positive flow values.

For flows recorded on December 20, 2002, the logarithm of the likelihood function is maximized at f(**x,** λ) = 4664 when λ = -0.61 (see Figure C2). Using the procedure, described in the thesis, leak analysis was carried out for the transformed data.

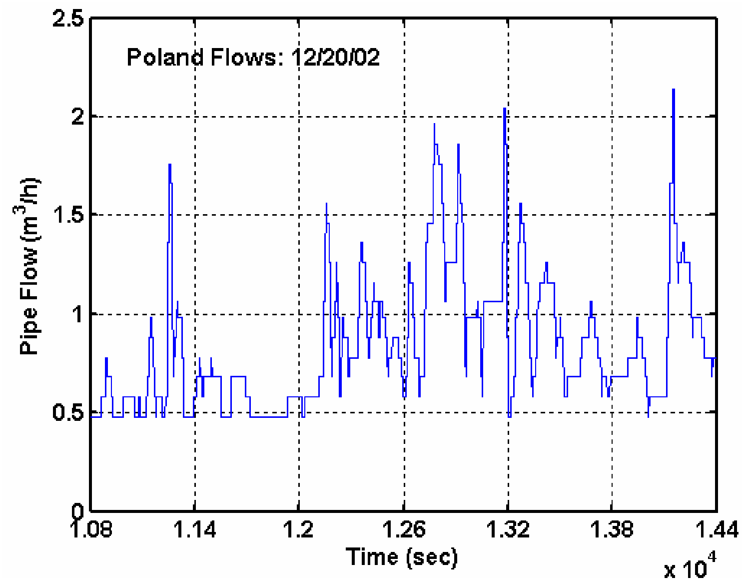
$\mu = 0.6127$ $\sigma = 0.4136$ $\varepsilon_{L1} = -1.4813$

Estimated leakage rate:

$Q_L' = 0.6127 + (-1.4813)\ (0.4136) = 0.00$

$$x(-0.61) = \frac{\left(x^{-0.61} - 1\right)}{-0.61} + 0.922 = 0.00$$
$$Hence, Q_L = 0.48 m^3 / h$$

The result agrees with the general observation made on page 78 of Appendix C. In the standardized plots of Figure C4, the observed curve follows a "saw-tooth" pattern on the standardized $S^{'}[Q^{*}]$ curve. The approximately horizontal lines of the observed curve correspond to discrete vertical spikes on the normal probability plot of transformed flows (seen from the linear part of normal probability plot in Figure C2). When the flow is constant for a certain period of time, the standard deviation will vary only slightly for each increment in the truncation level.



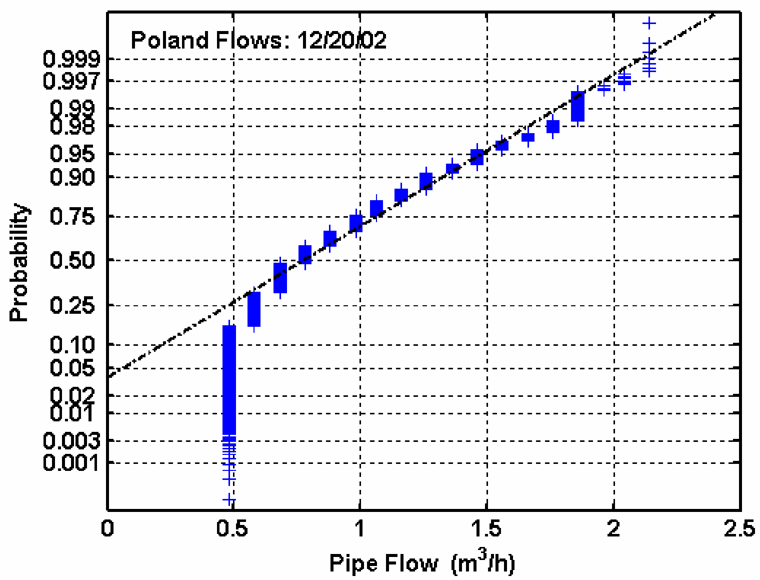


Figure C1 Poland flows (original data), December 20, 2002

Time series plot (top) and

Normal probability plot (bottom) of flows

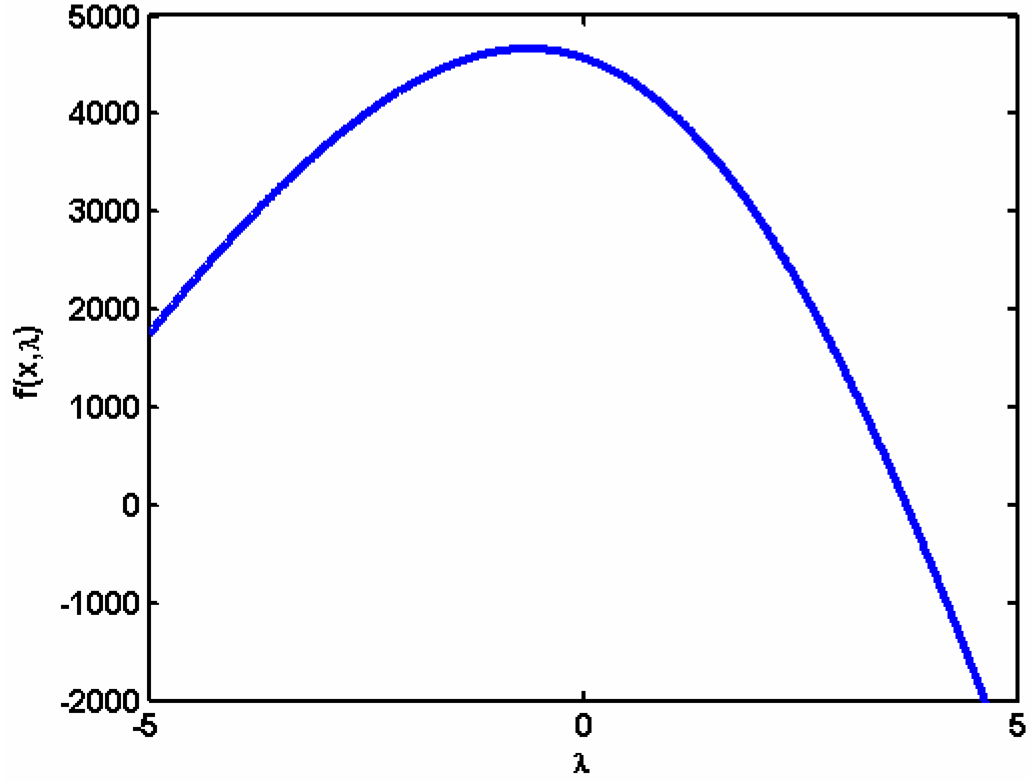
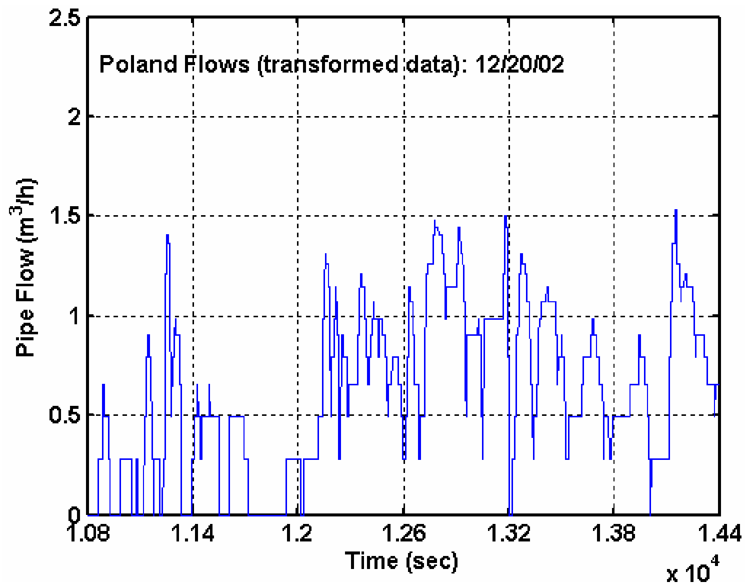


Figure C2 Plots of log likelihood function f(x, λ) versus λ



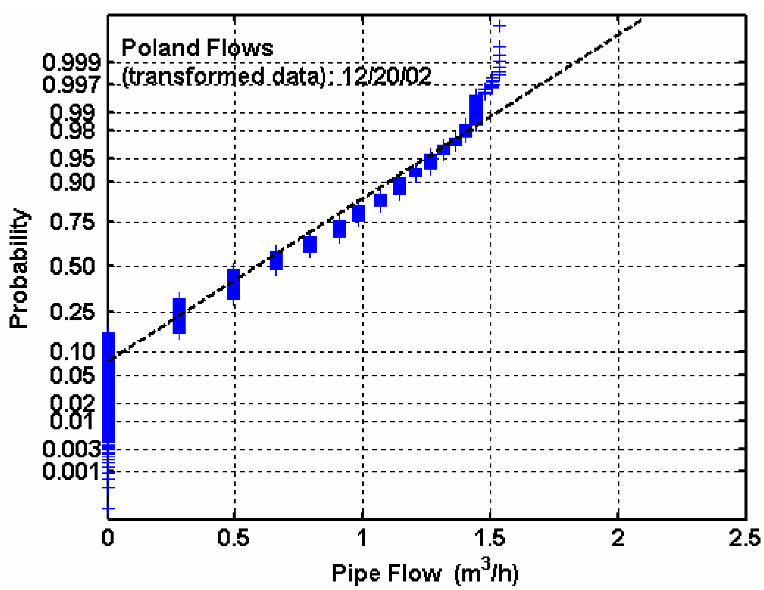


Figure C3 Poland flows (transformed data), December 20, 2002

Time series plot (top) and

Normal probability plot (bottom) of flows

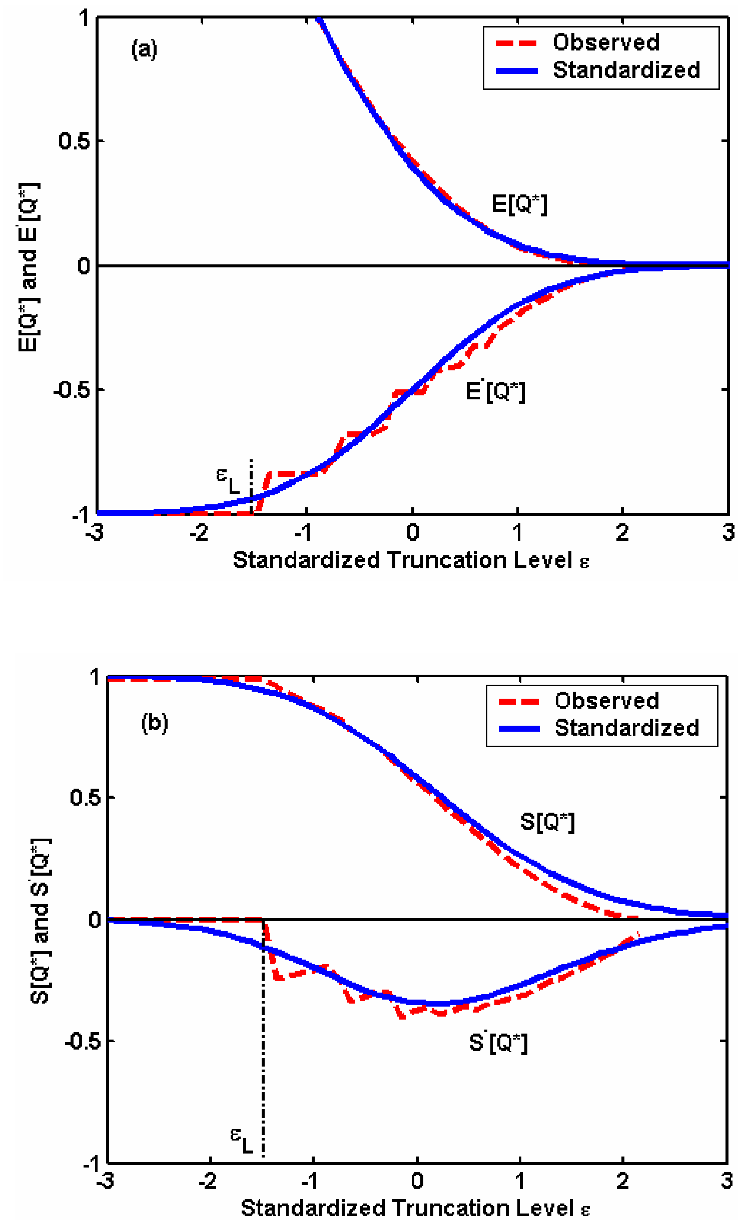


Figure C4 Observed and standardized curves for transformed data

Poland flows, December 20, 2002

APPENDIX D

C++ AND MATLAB CODES DEVELOPED FOR LEAK ANALYSIS

//CODE # 1

//THIS PROGRAM SIMULATES CONSTANT LEAKAGE RATES AND

//COMPUTES SEQUENTIAL STATISTICS OF TRUNCATED PIPE FLOWS

```
#include<iostream.h>

#include<fstream>

#include<stdio.h>

#include<stdlib.h>

#include<iomanip.h>

#include<math.h>

using namespace std;

const int T=86401;

const int P=2;


int main(int argc, char* argv[])

{

    char buffer[256];

   // ARRAY TO READ FLOWS TO A FILE

   int count,t;

   // COUNTERS
```

```
float **buffer1;
// ARRAY FOR INTEGRATING FLOWS
double **Q;
// ARRAY FOR STATISTICS OF TRUNCTATED FLOWS
float *A;
// A TEMPORARY ARRAY
float B[T][P];
// ARRAY FOR COPYING INTEGRATED FLOWS
float intn;
// COUNTER
float Qmax;
// VARIABLE FOR MAXIMUM FLOW
double s;
// COUNTER
 int p,t1,t2,time1, time2, rows;
 // COUNTERS
double temp1, temp2;
// TEMPORARY VARIABLES
long int temp3, temp4;
// TEMPORARY VARIABLES
cout << "\nInput the start and end time of detection in seconds" << endl;
cin >> t1 >> t2;
time1=t2-t1+2;
```

```
time2 = t2 - t1;

ifstream infile,infile2;

infile.open ("C:\\WINNT\\Profiles\\gayatri\\Desktop\\oct_16.txt" );

count=0;

if (! infile.is_open())

/ READS FLOW DATA TO A FILE

 cout << "Error opening file"; exit (1); }

while (! infile.eof() )

{

        infile.getline (buffer,100);

        count++;

 }

infile.close();

float k=0;

buffer1 = (float**) malloc(sizeof(float)*count);

for(int i=0;i<count-1;i++)

buffer1[i] = (float*)malloc(sizeof(float)*7);


infile2.open ("C:\\WINNT\\Profiles\\gayatri\\Desktop\\oct_16.txt" );

hile(!infile2.fail())

/ WRITES FLOW DATA TO A FILE

{

        for(int i=0;i<count-1;i++)
```

```
        for(int j=0;j<6;j++)
        infile2>>buffer1[i][j];
}

infile.close();
infile2.close();

if (buffer1==NULL) exit (1);
// COMPUTES END TIME OF EACH PULSE
for( i=0;i<count-1;i++)
        buffer1[i][6]=buffer1[i][2]+buffer1[i][3];

for(t=0;t<T;t++)
{
        for(int j=0;j<2;j++)
        {
                B[t][j]=0;
        }
}
i=0;
intn=0.0;

for(t=0;t<T;t++)
```

```
// INTEGRATES PULSES
{
        intn=0;
        while(i<count-1)
        {
                if(t>=buffer1[i][2] && t<=buffer1[i][6])
                {
                        intn = intn + buffer1[i][4];
                }
                i++;
        }
        i=0;
        B[t][0]=intn;
        B[t][1]=intn+1.0;
}
FILE *fp;
fp= fopen("oct_flows_C.txt", "a");


for(t=t1;t<t2;t++)
// WRITES INTEGRATED FLOWS TO A FILE
{
        fprintf(fp,"%d\t %2.4f\t %2.4f\n",t+1,B[t][0],B[t][1]);
}
```

```
fclose(fp);

double qTemp=0;

for(t=t1;t<t2;t++)

{

        if(qTemp<B[t][1])

        qTemp = B[t][1];

}


Qmax = qTemp;

rows = Qmax/0.05;

rows = rows + 1;

float max, min;


A = (float*) malloc(sizeof(float)*time1);

// INITIALIZES DYNAMIC TEMP ARRAY

Q=(double**)malloc(sizeof(double)*rows);

// INITIALIZES DYNAMIC ARRAY FOR

for(p=0;p<rows;p++)

// FLOW STATISTICS

Q[p] = (double*)malloc(sizeof(double)*11);


for(p=0;p<rows;p++)

{
```

```
for(int j=0;j<11;j++)
        {
                Q[p][j]=0;
        }
}
for(t=0;t<time1;t++)
          A[t]=0;
for(t=t1;t<=t2;t++)
A[t-t1]=B[t][1];


s=0.0;
for(p=0;p<rows;p++)
// COMPUTES STATISTICS OF TRUNCATED FLOWS
{
        temp1=0.0;
        temp2=0.0;
        temp3=0.0;
        temp4=0.0;
        max = 0.0;
        min = 0.0;
        for(t=t1;t<t2;t++)
        {
                A[t-t1]=B[t][1]-s;
```

```
        if(A[t-t1]<0){
          A[t-t1]=0;
          temp4+=1;}
        temp1+= A[t-t1]*1;
        temp2+= A[t-t1]*A[t-t1]*1;
        temp3+= 1;
        A[t-t1]=0;
}
Q[p][0] = (temp1/temp3);
// MEAN
Q[p][1] = (temp2/temp3);
// MEAN SQUARE
Q[p][2]= (temp2/temp3)-((temp1/temp3)*(temp1/temp3));
if(Q[p][2]<0)
{ Q[p][3]=0;}
else
{ Q[p][3] = Q[p][2];}
 // VARIANCE
Q[p][4] = sqrt(Q[p][3]);
 // STANDARD DEVIATION
Q[p][5] = Q[p][4]/Q[p][0];
Q[p][6] = temp4;
Q[p][7] = temp3;
```

```
        Q[p][8] = Q[p][6]/Q[p][7];
        // STAGNATION PROBABILITY
        s=s+0.05;
    }


    FILE *fc;
    fc= fopen("oct_stats_C.txt", "a");
    s=0.0;
    for(p=0;p<rows;p++)
    {
        fprintf(fc,"%2.4f\t %2.4f\t %2.4f\t %2.4f\t %2.4f\t %2.4f\t %2.4f\t %2.4f\n",
        s,Q[p][0],Q[p][3],Q[p][4],Q[p][5],Q[p][8]);
        s=s+0.05;
    }
    fclose(fc);
    return 0;
}
```

```
// CODE # 2
// THIS PROGRAM GENERATES SIMULATED VARIABLE LEAKAGE RATES AND
COMPUTES
// SEQUENTIAL FLOW STATISTICS FOR TRUNCATED FLOWS

    #include<iostream.h>
    #include<fstream>
    #include<stdio.h>
    #include<stdlib.h>
    #include<iomanip.h>
    #include<math.h>
    #include<time.h>
    #include<conio.h>
    using namespace std;

    const int T=86401;
    const int P=2;

    int main(int argc, char* argv[])
    {
        char buffer[256];
        // ARRAY TO READ FLOWS TO A FILE
        float B[T][P];
```

```
// ARRAY FOR INTEGRATED FLOWS
float **buffer1;
// ARRAY TO WRITE FLOWS TO A FILE
double **Q;
// ARRAY FOR STATISTICS OF TRUNCATED FLOWS
double *A,**scr;
// TEMPORARY ARRAY
float intn;
// COUNTER
float Fmax, Qmax;
// MAXIMUM FLOWS
float r1,r2,x1,x2,w,S,R,s;
// VARAIBLES FOR SRANDARD NORMAL RANDOM NUMBERS
int count,t,t1,t2,time1,rows,p,time2;
// COUNTERS
double temp1,temp2;
// TEMPORARY VARIABLES
int temp3,temp4;
// TEMPORARY VARIABLES


cout <<"\nEnter the Covariance of the leakage rate(S): ";
cin >> S;
cout <<"\nEnter the autocorrelation coefficient for leakage rate(R): ";
```

```
cin >> R;
cout << endl << endl;

cout << "\nInput the start and end time of detection in seconds" << endl;
cin >> t1 >> t2;
time1=t2-t1+2;
time2=t2-t1;

ifstream infile,infile2;
infile.open ("C:\\WINNT\\Profiles\\gayatri\\Desktop\\2000homes2-2hrs.txt" );
count=0;

if (! infile.is_open())
// READS FLOW DATA TO A FILE
{ cout << "Error opening file"; exit (1); }

while (! infile.eof() )
{
        infile.getline (buffer,100);
        count++;
}
infile.close();
float k=0;
```

```
buffer1 = (float**) malloc(sizeof(float)*count);
for(int i=0;i<count-1;i++)
buffer1[i] = (float*)malloc(sizeof(float)*7);

infile2.open ("C:\\WINNT\\Profiles\\gayatri\\Desktop\\2000homes2-2hrs.txt" );
while(!infile2.fail())
// WRITES FLOW DATA TO A FILE
{
	for(int i=0;i<count-1;i++)
		for(int j=0;j<7;j++)
	infile2>>buffer1[i][j];
}
infile.close();
infile2.close();

if (buffer1==NULL) exit (1);
// COMPUTES END TIME OF EACH PULSE
for( i=0;i<count-1;i++)
	buffer1[i][7]=buffer1[i][3]+buffer1[i][4];

for(t=0;t<T;t++)
{
	for(int j=0;j<2;j++)
```

```
		{
			B[t][j]=0;
		}
}


i=0;
for(t=0;t<T;t++)
// INTEGRATES PULSES
{
		intn=0;
		while(i<count-1)
		{
			if(t>=buffer1[i][3] && t<=buffer1[i][7])
			{
				intn = intn + buffer1[i][5];
			}
			i++;
		}
		i=0;
B[t][0]=intn;
}
double flowTemp=0.0;
for(t=0;t<T;t++)
```

```
// FINDS MAXIMUM FLOW WITHIN THE OBSERVATION WINDOW
{
        if(flowTemp<B[t][0])
                flowTemp=B[t][0];
}
Fmax = flowTemp;
printf("\nFmax=%2.4f\n", Fmax);

for(t=0;t<T;t++)
// COMPUTES MEAN LEAKAGE RATES FOR EACH FLOW
{
        B[t][1] = 2.5-(0.5*B[t][0]/Fmax);
}

Q=(double**)malloc(sizeof(double)*T);
for(t=0;t<T;t++)
        Q[t] = (double*)malloc(sizeof(double)*3);

for(t=0;t<T;t++)
{
        for(int j=0;j<3;j++)
        {
                Q[t][j]=0;
```

```
        }
}

srand(time(0));
for(t=0;t<T;t++)
// GENERATES STANDARD NORMAL RANDOM NUMBERS N~(0,1)
{
        x1=0.0;
        x2=0.0;
        w=0.0;

        r1 = (float)(rand()%RAND_MAX)/RAND_MAX;
        r2 = (float)(rand()%RAND_MAX)/RAND_MAX;

        x1 = 2*r1-1;
        x2 = 2*r2-1;
         w = x1*x1 + x2*x2;

        if(w <= 1){
                w = sqrt((-2.*log(w))/w);
                x1 = x1*w;}
                 Q[t][0] = x1;
         }
```

```
Q[0][2] = 2.5;

for(t=0;t<T;t++)

// GENERATES VARIABLE LEAKS FOR COMPUTES FLOWS

{

// WITH VARIABLE LEAKAGE

        if(t>0)

        {

                Q[t][1] = B[t][1] + R*(Q[t-1][1]-B[t][1]) + Q[t][0]*S*B[t][1]*(sqrt(1-
                (R*R)));

                Q[t][2] = B[t][0] + Q[t][1];

        }

}

cout<<Q[0][2];


FILE *flows;

// WRITES VARIABLE LEAKS AND CORRESPONDING FLOWS TO A //FILE

flows = fopen("2000homesflows2-V-2hrs.txt", "a");

for(t=t1;t<t2;t++)

{

        fprintf(flows,"%d\t %2.4f\t %2.4f\t",t+1,B[t][0],B[t][1]);

        fprintf(flows,"%2.6f\t %2.4f\t %2.4f\n",Q[t][0],Q[t][1],Q[t][2]);


}
```

```
fclose(flows);


double QTemp=0.0;

for(t=t1;t<t2;t++)

// FINDS MAXIMUM FLOW VALUE WITHIN THE OBSERVATION WINDOW

{

	if(QTemp<Q[t][2])

		QTemp=Q[t][2];

}

Qmax = QTemp;

printf("\nQmax=%2.4f\n\n",Qmax);


rows = Qmax/0.05;

rows=rows+1;

cout << rows;


A = (double*) malloc(sizeof(double)*time1);

// INITIALIZES DYNAMIC TEMPORARY ARRAYS


scr=(double**)malloc(sizeof(double)*rows);

// INITIALIZES DYNAMIC ARRAY FOR FLOW STATISTICS

	for(p=0;p<rows;p++)

		scr[p] = (double*)malloc(sizeof(double)*11);
```

```
for(p=0;p<rows;p++)

{

            for(int j=0;j<=11;j++)

            {

                        scr[p][j]=0;

            }

}

for(t=0;t<time1;t++)

            A[t]=0;

for(t=t1;t<t2;t++)

A[t-t1]=Q[t][2];

s=0.0;

for(p=0;p<rows;p++)

// COMPUTES STATISTICS OF TRUNCATED FLOWS

{

            temp1=0.0;

            temp2=0.0;

            temp3=0.0;

            temp4=0.0;

            max = 0.0;

            min = 0.0;

            for(t=t1;t<t2;t++)

            {
```

```
        A[t-t1]=Q[t][2]-s;


        if(A[t-t1]<0){
          A[t-t1]=0;
          temp4+=1;}


        temp1+= A[t-t1];
        temp2+= A[t-t1]*A[t-t1];
        temp3+= 1;
}
scr[p][0] = (temp1/temp3);
// MEAN
scr[p][1] = (temp2/temp3);
// MEAN SQUARE
scr[p][2]       = (temp2/temp3)-((temp1/temp3)*(temp1/temp3));
if(scr[p][2]<0)
        { scr[p][3]=0;}
else
        { scr[p][3] = scr[p][2];}
// VARIANCE
scr[p][4] = sqrt(scr[p][3]);
// STANDARD DEVIATION
scr[p][5] = scr[p][4]/scr[p][0];
```

```
        scr[p][6] = temp4;
        scr[p][7] = temp3;
        scr[p][8] = scr[p][6]/scr[p][7];
        // STAGNATION PROBABILITY
        scr[p][9] = max;
        scr[p][10] = min;
        s=s+0.05;
}
FILE *fp;
fp= fopen("2000homesstats2-V-2hrs.txt", "a");
s=0.0;
for(p=0;p<rows;p++)
// WRITES STATISTICS OF TRUNCATED FLOWS TO A FILE
{
        fprintf(fp,"%2.4f\t %2.4f\t %2.4f\t", s,scr[p][0],scr[p][3]);
        fprintf(fp,"%2.4f\t %2.4f\t %2.4f\t %2.4f\t %2.4f\n",scr[p][4],scr[p][5],scr[p][8],
        scr[p][9], scr[p][10]);
        s=s+0.05;
}
printf("\n\nMaximum Flow = %2.4f\n Max Flow in the obervation window=%2.4f\n
rows=%d\n\n", Fmax,Qmax,rows);
fclose(fp);
return 0; }
```

```
% CODE # 3
% THIS PROGRAM ESTIMATES PARAMETERS OF PARENT NORMAL
DISTRIBUTION
% FOR THE CONSTANT LEAKAGE CASE USING LEAST SQUARES REGRESSION

function parameters = analAlgorithm(cutoffs)
% PASSES REGRESSION PARAMETERS
Y = load('N:\matlab\may_01_1.txt');
% LOADS TEXT FILE OF FLOWS
Y(:,2) = Y(:,2).*3.785;
% CONVERTS PIPE FLOWS WITHOUT LEAKAGE INTO L/MIN UNITS
Y(:,3) = Y(:,3).*3.785;
% CONVERTS PIPE FLOWS WITH CONSTANT LEAKAGE INTO L/MIN UNITS

X = [];
% INITIALIZES AVECTOR FOR FLOWS WITH LEAKS
sortedFlows1 = [];
% INITIALIZES VECTOR OF SORTED FLOWS
t = [];
% INITIALIZES VECTOR TO CALCULATE STANDARD NORMAL VARIATES
a = [];
b = [];
z = [];
```

```
Z = [];
% INITIALIZES VECTOR FOR STANDARD NORMAL VARIATES
parameters = [];
% INITIALIZES VECTOR FOR REGRESSION PARAMETERS


X = Y(:,3);
% COPIES FLOWS WITH LEAKS INTO VECTOR 'X'
X=sort(X);
% SORTS FLOWS IN ACSENDING ORDER
n=length(X);
% COMPUTES LENGTH OF SORTED FLOWS


for i=1:n
% THIS LOOP RE-ARRANGES FLOWS IN ASCENDING ORDER
   sortedFlows1(i,1)=X(n-i+1);
end


for i=1:cutoffs
% ASSIGNS RANK TO EACH FLOW
   sortedFlows1(i,2) = i;
end
pe = (sortedFlows1(:,2)-0.375)./(cutoffs+0.25);
% COMPUTES p^TH QUANTILE OF STANDARD NORMAL DISTRIBUTION
```

```
for i=1:n
% COMPUTES THE STANDARD NORMAL VARIATES
   if(pe(i) < 0.5)
       t(i) = sqrt(-2*log(pe(i)));
       a(i) = 2.515517 + 0.802853*t(i) + 0.010328*t(i)^2;
       b(i) = 1 + 1.432788*t(i) + 0.189269*t(i)^2 + 0.001308*t(i)^3;
       z(i)= t(i) - (a(i)/b(i));
   elseif(pe(i) > 0.5)
       pn(i)=1-pe(i);
       t(i) = sqrt(-2*log(pn(i)));
       a(i) = 2.515517 + 0.802853*t(i) + 0.010328*t(i)^2;
       b(i) = 1 + 1.4327*t(i) + 0.189269*t(i)^2 + 0.001308*t(i)^3;
       z(i) = -(t(i)-(a(i)/b(i)));
   else
      if(pe(i)==0.5)
           z(i)=0;
   end
end
Z=[Z;z(i)];
end
k1=size(Z);


[parameters S] = polyfit(Z(1:cutoffs,1),sortedFlows1(1:cutoffs,1),1);
```

```
sigma = parameters(1);
% SLOPE OF LINEAR REGRESSION
mu = parameters(2);
% INTERCEPT OF LINEAR REGRESSION

figure(1)
% PLOTS PIPE FLOWS (WITHOUT LEAKAGE) VERSUS TIME
plot(Y(:,1),Y(:,2))
xlabel('Time (sec)', 'FontSize', 12.0, 'FontWeight', 'bold')
ylabel('Pipe Flow (L/min)', 'FontSize', 12.0, 'FontWeight', 'bold')
text(50,56, '(a) No leakage', 'FontSize', 12.0, 'FontWeight', 'bold')
grid on;
axis([0 3600 0 70])
set(gca, 'xtick',[0, 600, 1200, 1800, 2400, 3000, 3600], 'FontSize', 12.0, 'FontWeight', 'bold')
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)

figure(2)
% PLOTS PIPE FLOWS (WITH CONSTANT LEAKAGE) VERSUS TIME
plot(Y(:,1),Y(:,3))
xlabel('Time (sec)', 'FontSize', 12.0, 'FontWeight', 'bold')
ylabel('Pipe Flow  (L/min)', 'FontSize', 12.0, 'FontWeight', 'bold')
text(50,56, '(b) Constant leakage (Q_L = 3.79 L/min)', 'FontSize', 12.0, 'FontWeight', 'bold')
axis([0 3600 0 70])
```

```
set(gca, 'xtick',[0, 600, 1200, 1800, 2400, 3000, 3600], 'FontSize', 12.0, 'FontWeight', 'bold')
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
grid on;

figure(3)
% PLOTS NORMAL PROBABILITY PLOTS OF PIPE FLOWS WITHOUT LEAKAGE
normplot(Y(:,2))
xlabel('Pipe Flow  (L/min)', 'FontSize', 12.0, 'FontWeight', 'bold')
ylabel('Probability', 'FontSize', 12.0, 'Fontweight', 'bold')
text(2,3.1, '(a) No leakage', 'FontSize', 12.0, 'FontWeight', 'bold')
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
axis([0 80 -3.81 3.81])

figure(4)
% PLOTS NORMAL PROBABILITY PLOTS OF PIPE FLOWS WITH CONSTANT
LEAKAGE
normplot(Y(:,3))
xlabel('Pipe Flow  (L/min)', 'FontSize', 12.0, 'FontWeight', 'bold')
ylabel('Probability', 'FontSize', 12.0, 'Fontweight', 'bold')
text(2,3.1, '(b) Constant leakage (Q_L = 3.79 L/min)', 'FontSize', 12.0, 'FontWeight', 'bold')
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
axis([0 80 -3.81 3.81])
```

```
path = ('N:\matlab\PDF.txt');

fid1 = fopen('N:\matlab\test7.txt','w');

% WRITES STANDARD NORMAL VARIATES INTO A TEXT FILE.

for i=1:cutoffs

   fprintf(fid1,'12.4f\n', Z(i,1));

end

fclose(fid1);


path = ('N:\matlab\FLOWS.txt');

fid2 = fopen('N:\matlab\test8.txt','w');

% WRITES RANKED FLOWS INTO A TEXT FILE.

for i=1:cutoffs

   fprintf(fid2,'12.4f\n', sortedFlows1(i,1));

end

fclose(fid2);
```

% CODE # 4

% THIS PROGRAM PASSES THE CUTOFF VALUE TO CODE # 3 AND ESTIMATES THE LINEAR REGRESSION PARAMETERS

```
function StdData = driftValue

cutoffs = 839;
% CUTOFF VALUE

mstd = analAlgorithm(cutoffs);
% PASSES CUTOFF VALUE TO ESTIMATE
slope = mstd(1);
% REGRESSION PARAMETERS
intercept = mstd(2);
StdData = [slope intercept];
% SLOPE AND INTERCEPT FOR LINEAR REGRESSION
```

% CODE # 5

% THIS PROGRAM PLOTS STANDARDIZED CURVES FOR MOMENTS

% (MEAN AND STANDARD DEVIATION) AND THEIR DERIVATIVES OF PIPE FLOWS

```
function theoreticalValues = moments


Q=load ('N:\matlab\apr_VFlows.txt');
% LOADS EITHER PIPE FLOWS WITH CONSTANT
Q=Q(:,6).*3.785;
% LEAKAGE OR VARAIBLE LEAKAGE
Qmax = max(Q);
% EVALUATES MAXIMUM FLOW


Ql=[];
% INITIALIZES VECTOR FOR TRUNCATION LEVEL
Pl=[];
%standardizedDrift=[];
% INITIALIZES VECTOR FOR STANDARDIZED TRUNCATION VALUES
Ql = -100:0.05:Qmax+10;
k=length(Ql);


constants = driftValue;
% TAKES REGRESSION PARAMETERS
```

```
slope = constants(1);

intercept = constants(2);


epsilon = (Ql - intercept)/slope;
```

% INITIALIZES VECTOR FOR STANDARDIZED TRUNCATION VALUES

```
normalProb = [];
```

% INITIALIZES VECTOR FOR PDF

```
PHIprob = [];

normalCdf = [];
```

% INITIALIZES VECTOR FOR CDF

```
normalCdf1 = [];

ABS = [];

CDFfunction1 = [];

Mean=[];

squareMean=[];

Variance=[];

StdDev = [];

StagProb = [];

standardValues = [];

slopeMean = [];

slopeStddev = [];

derivative = [];
```

```
denominator = [];

slopeProb = [];


k=1/sqrt(2*pi);

ABS = abs(epsilon);


normalProb = k*exp(-(epsilon.^2)/2);

% COMPUTES PDF VALUES

n=length(normalProb);

% FINDS LENGTH OF PDF VECTOR


CDFfunction1 = 0.5*(1 + 0.0498673470.*ABS+ 0.0211410061*ABS.^2 + ...

        0.0032776263*ABS.^3 + 0.0000380036*ABS.^4 + ...

        0.0000488906*ABS.^5 + 0.0000053830*ABS.^6).^-16;


for i=1:n

% COMPUTES CDF VALUES

  if(epsilon(i)>=0)

    PHIprob(i)=1-CDFfunction1(i);

  else

    PHIprob(i)=CDFfunction1(i);

  end

    normalCdf = [normalCdf;PHIprob(i)];
```

```
end

normalCdf1 = 1-normalCdf;

normalCdf=normalCdf';


Mean = normalProb - epsilon.*(normalCdf1');

% EVALUATES MEAN OF PIPE FLOWS

squareMean = -epsilon.*normalProb + (1 + epsilon.^2).*(normalCdf1');

Variance = squareMean - Mean.^2;

% EVALUATES VARIANCE OF PIPE FLOWS

StdDev = sqrt(Variance);

% EVALUATES STANDARD DEVIATION OF PIPE FLOWS

StagProb = normalCdf;


slopeMean = -normalCdf1';

% EVALUATES DERIVATIVE OF MEAN OF TRUNCATED PIPE FLOWS


denominator = sqrt(Variance);

derivative = normalCdf.*(epsilon - normalProb - epsilon.*normalCdf);

slopeStddev = derivative./denominator;

% EVALUATES DERIVATIVE OF STANDARD DEVIATION OF PIPE FLOWS


 figure(2);

% PLOTS STANDARDIZED CURVES FOR MEAN AND STANDARD DEVIATION
```

```
plot(epsilon, StdDev,'LineWidth', 3)
%OF PIPE FLOWS
hold on
plot(0.5, 0.4308, 'LineStyle', 'none', 'Marker', '^', 'MarkerSize', 9.0, 'MarkerEdgeColor', 'black',
'MarkerFaceColor', 'black');
hold on
plot(epsilon, Mean, 'LineWidth', 3, 'color', 'red')
hold on
plot(0.5, 0.1950, 'LineStyle', 'none', 'Marker', 'd', 'MarkerSize', 9.0, 'MarkerEdgeColor', 'black',
'MarkerFaceColor', 'black');
xlabel('Standardized Truncation Level \epsilon','FontSize', 12.0, 'FontWeight', 'bold')
ylabel('E[Q*] and S[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
text(-2, 1.1,'S[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(-2, 2.25, 'E[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(-2.5, 2.75, '(a) Moments', 'FontSize', 12.0, 'FontWeight', 'bold')
axis([-3 3 0 3])
grid on
hold off

figure(3)
% PLOTS STANDARDIZED CURVES FOR DERIVATIVES OF MEAN
plot(epsilon, slopeMean, 'LineWidth', 3, 'color', 'red')
```

```
% AND STANDARD DEVIATION OF PIPE FLOWS
hold on
plot(-0.72, -0.73, 'LineStyle', 'none', 'Marker', 'd', 'MarkerSize', 8.0)
hold on
plot(epsilon, slopeStddev, 'LineWidth', 3)
hold on
plot(-0.72, -0.25, 'LineStyle', 'none', 'Marker', 'd', 'MarkerSize', 8.0)
xlabel('Standardized Truncation Level \epsilon', 'FontSize', 12.0, 'FontWeight', 'bold')
ylabel('E^\prime[Q*] and S^\prime[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(-2.0,-0.15, 'S^\prime[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(-0.6,-0.45, 'E^\prime[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(0, -0.9, '(b) Derivatives of moments', 'FontSize', 12.0, 'FontWeight', 'bold')
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
axis([-3 3 -1 0])
grid on
hold off


theoreticalValues = [epsilon' Mean' StdDev' StagProb' slopeMean' slopeStddev'];
% STANDARDIZED VALUES OF
%MOMENTS AND THEIR DERIVATIVES
```

```
% CODE # 6
% THIS PROGRAM PLOTS CURVES FOR OBSERVED SEQUENTIAL STATISTICS
OF TRUNCATED
% PIPE FLOWS AND SUPERIMPOSES THEM ON TO STANDARDIZED CURVES.

file = 'N:\matlab\apr_Vstats.txt';
Stats = load(file);
% LOADS OBSERVED VALUES OF OBSERVED MEAN AND
% STANDARD DEVIATION OF TRUNCATED PIPE FLOWS
Censorlevel = [];
% INITIALIZES VECTOR FOR TRUNCATION LEVEL
MeanStats = [];
% INITIALIZES VECTOR FOR MEAN OF TRUNCATED PIPE FLOWS
StdevStats = [];
% INITIALIZES VECTOR FOR MEAN OF TRUNCATED PIPE FLOWS
%standardValues = [];
slopeMean = [];
% EVALUATES DERIVATIVES OF MEAN OF TRUNCATED PIPE FLOWS
slopeDev = [];
% EVALUATES DERIVATIVES OF STANDARD DEVIATION
%OF TRUNCATED PIPE FLOWS
A = [];
Censorlevel1 = [];
```

```
Censorlevel1 = -20:0.05:-0.05;

p=length(Censorlevel1);

a = Stats(1,2) + 20;

b = Stats(1,2) + 0.05;

MeanStats1 = a:-0.05:b;

StdevStats1(1:p) = Stats(1,4);

ProbStats1(1:p) = Stats(1,6);


Censorlevel = [Censorlevel1'.*3.785; Stats(:,1).*3.785];

MeanStats = [MeanStats1'.*3.785; Stats(:,2).*3.785];

StdevStats = [StdevStats1'.*3.785; Stats(:,4).*3.785];

length(Censorlevel);

constants = driftValue;

slope = constants(1);

intercept = constants(2);


StandardizedMean = MeanStats./slope;

% STANDARDIZES THE OBSERVED MEAN VALUES

StandardizedStdev = StdevStats./slope;

% STANDARDIZES THE OBSERVED STANDARD DEVIATION

stdCensorlevel = (Censorlevel - intercept)./slope;

% STANDARDIZES TRUNCATION LEVELS

B=[];
```

```
n = length(MeanStats);
for i=2:n
% EVALUATES SLOPES OF MEAN AND STANDARD DEVIATION OF
   k1 = (MeanStats(i)-MeanStats(i-1))/(0.05.*3.785);
   % TRUNCATED PIPE FLOWS
   k2 = (StdevStats(i)-StdevStats(i-1))/(0.05.*3.785);
   slopeMean = [slopeMean; k1];
slopeDev = [slopeDev; k2];
end

slopeMean(1) = slopeMean(2);
slopeMean = [slopeMean(1);slopeMean];

slopeDev(1) = slopeDev(2);
slopeDev = [slopeDev(1);slopeDev];
constants
B=[stdCensorlevel StandardizedMean slopeMean StandardizedStdev slopeDev]
C1 = -3:0.25:3;
C2(1:25) = 0.0;
standardValues = moments;

figure(7);
% PLOTS MEAN AND ITS DERIVATIVE OF TRUNCATED PIPE FLOWS AND
```

```
plot(stdCensorlevel, StandardizedMean, 'linestyle','--','color','r','linewidth',3.0);
hold on;
% SUPERIMPOSES ONTO STANDARDIZED CURVES
plot(standardValues(:,1) , standardValues(:,2), 'linewidth', 3.0);
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
legend('Observed', 'Standardized', 1);
hold on;


plot(stdCensorlevel, slopeMean, 'linestyle','--','color','r','linewidth', 3.0);
hold on;
plot(standardValues(:,1) , standardValues(:,5), 'linewidth', 3.0);
xlabel('Standardized Truncation Level \epsilon', 'FontSize', 12.0, 'FontWeight', 'bold')
%ylabel('E[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
ylabel('E[Q*] and E^\prime[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(-1.5, 0.75, '(a)', 'FontSize', 12.0, 'FontWeight', 'bold')
text(0.75, 0.25, 'E[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold');
text(0.5, -0.5, 'E^\prime[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold');
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
axis([-3 3 -1 1])


hold on;
plot(C1',C2', 'color', 'black', 'LineWidth', 2.0);
hold off;
```

```
figure(8);
%PLOTS STANDARD DEVIATION AND ITS DERIVATIVE OF TRUNCATED
plot(stdCensorlevel, StandardizedStdev, 'linestyle','--','color','r','linewidth',3.0)
hold on
%PIPE FLOWS AND SUPERIMPOSES ONTO STANDARDIZED CURVES
plot(standardValues(:,1), standardValues(:,3), 'linewidth', 3.0)
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
legend('Observed', 'Standardized', 1)
hold on;


plot(stdCensorlevel, slopeDev, 'linestyle','--','color','r','linewidth',3.0)
hold on
plot(standardValues(:,1), standardValues(:,6), 'linewidth', 3.0)
xlabel('Standardized Truncation Level \epsilon', 'FontSize', 12.0, 'FontWeight', 'bold')
%ylabel('S[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
ylabel('S[Q*] and S^\prime[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(-2.75, 0.75, '(b)', 'FontSize', 12.0, 'FontWeight', 'bold')
text(1.0, 0.4, 'S[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
text(0.5, -0.5, 'S^\prime[Q*]', 'FontSize', 12.0, 'FontWeight', 'bold')
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
axis([-3 3 -1 1])
```

```
hold on;
plot(C1',C2', 'color', 'black', 'LineWidth', 2.0);
hold off;
```

constants

```
% CODE # 7

% THIS PROGRAM PLOTS BOX AND WHISKERS PLOTS FOR SIMULATED

LEAKAGE RATES AND

% COMPARES WITH ESTIMATED LEAKAGE RATES


function leakSummary

X1 = load('N:\matlab\apr_VFlows.txt');

% STATEMENTS 6 - 24 LOAD SIMULATED LEAKAGE

size(X1);

% RATES FOR VARIOUS SAMPLE DATA SETS

X2 = load('N:\matlab\may_VFlows.txt');

X3 = load('N:\matlab\jun_VFlows.txt');

X4 = load('N:\matlab\jul_VFlows.txt');

X5 = load('N:\matlab\aug_VFlows.txt');

X6 = load('N:\matlab\sep_VFlows.txt');

X7 = load('N:\matlab\oct_VFlows.txt');

X8 = load('N:\matlab\200homesflows-0.75-V-2.txt');

X9 = load('N:\matlab\400homesflows-V.txt');

X10 = load('N:\matlab\500homesflows-V.txt');


X = [X1(:,5) X2(:,5) X3(:,5) X4(:,5) X5(:,5) X6(:,5) X7(:,5) X8(:,5) X9(:,5), X10(:,5)];

X = X.*3.785;

clear X1 X2 X3 X4 X5 X6 X7 X8 X9
```

```
Y = load('N:\matlab\estLeakRates.txt');
Y(:,2)=Y(:,2).*3.785;

figure(1)
% PLOTS BOX AND WHISKERS PLOTS FOR SIMULATED LEAKAGE
boxplot(X(:,:))
% RATES AND COMPARES WITH ESTIAMTED LEAKAGE RATES
hold on
z = plot(Y(:,1),Y(:,2),'Marker','o', 'MarkerSize', 6.0, 'color','black','linestyle', 'none');
axis([0 11 0 25.0])
set(gca,'YGrid','on')
set(gca, 'xticklabel',{'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'S-200', 'S-400' ,'S-500'});
xlabel('Data Set', 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0);
ylabel('Leakage Rate (L/min)', 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0);
text(0.25, 21, 'Apr through Oct are measured flows', 'FontSize', 12.0, 'FontWeight', 'bold',
'LineWidth', 2.0)
text(0.25, 19.5, 'S-200, 400, 500 are simulated flows', 'FontSize', 12.0, 'FontWeight', 'bold',
'LineWidth', 2.0)
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
hold off
```

```
% CODE # 8
% THE FOLLOWING CODE SEGMENT COMPARES SIMULATED AND
PREDICTED STAGNATIONS
A =load('N:\matlab\stagnation.txt);
K=load('N:\matlab\preStagnations.txt');
figure(1)
% PLOTS SIMULATED AND PREDICTED STAGNATION VALUES
boxplot(A(:,:))
hold on
z = plot(K(:,1),K(:,2),'Marker','o', 'MarkerSize', 6.0, 'color','black', 'linewidth', 2.0);
set(gca,'YGrid','on')
set(gca, 'xticklabel',{'0','10','20','30','40','50','60','70','80','90','100', '110', '120','130', '140', '150',
'160', '170', '180', '190', '200'}, 'FontSize', 12.0, 'FontWeight', 'bold');
xlabel('Time Step \Deltat (sec)');
ylabel('Stagnation P[0] (%)')
text(5.5, 28, 'Observation Window = 12 hrs', 'FontSize', 12.0, 'FontWeight', 'bold')
text(5.5, 26, 'Mean Pulse Intensity = 8.52 L/min', 'FontSize', 12.0, 'FontWeight', 'bold')
text(5.5, 24, 'Mean Pulse Duration = 0.75 min', 'FontSize', 12.0, 'FontWeight', 'bold')
legend('- Computed Stagnation','-o- Predicted Stagnation')
axis([1 20 0 40])
set(gca, 'FontSize', 12.0, 'FontWeight', 'bold', 'LineWidth', 2.0)
hold off
```