VIRGINIA WATER RESOURCES RESEARCH CENTER

OPTIMAL DESIGN-REHABILITATION STRATEGIES FOR RELIABLE WATER DISTRIBUTION SYSTEMS

SPECIAL REPORT

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
BLACKSBURG, VIRGINIA

2002

Funds for the support of this project were provided by the Virginia Water Resources Research Center and the U.S. Geological Survey under the grant 1434-HQ-96-GR-02703. The contents of this publication do not necessarily reflect the views or policies of the Virginia Water Resources Research Center or the U.S. Geological Survey. The mention of commercial products, trade names, or services does not constitute an endorsement or recommendation.

Copies of this publication are available on the Center's website:
www.vwrrc.vt.edu

Printed Copies are $10.00

Virginia Water Resources Research Center
10 Sandy Hall
Blacksburg, VA 24061
(540)231-5624
FAX: (540)231-6673
e-mail: water@vt.edu

Leonard A. Shabman, Director

OPTIMAL DESIGN-REHABILITATION STRATEGIES FOR RELIABLE WATER DISTRIBUTION SYSTEMS

G.V. Loganathan
H.D. Sherali
S. Park
S. Subramanian

Editor: Judy Poff

Virginia Water Resources Research Center
10 Sandy Hall, 0444
Blacksburg, Virginia 24061

May 2002

VWRRC Special Report SR-20-2002

ACKNOWLEDGEMENTS

The writers greatly appreciate the financial support provided by the U.S. Geological Survey under the grant 1434-HQ-96-GR-02703 which made this research possible. The writers thank Mrs. Inga Solberg, Drs. Leonard Shabman and Tamim Younos of the Virginia Water Resources Research Center for their help at the various stages of the project. In addition, the writers would like to thank the Blacksburg town engineers, R. Formica and K. Mattingly for providing the distribution system data, and Messrs. J. Stathis, and R. Totlani for the earlier testing of the ideas presented in this report. The writers are indebted to Drs. A.K. Deb, Y. Hasit, and Messrs. F. Grablutz, and J. Snyder of the Roy F. Weston. Inc., for sharing many valuable ideas.

Final report submitted to the US Geological Survey based on the completion of the research project funded under the grant 1434-HQ-96-GR-02703

TABLE OF CONTENTS

CHAPTER 1

EXECUTIVE SUMMARY

The condition of water distribution systems plays a crucial role in a society because of its strong correlation with the community health standards and the potential for future economic growth. Aging, deteriorating water systems have long been recognized as a critical national problem. Pipes have become heavily turberculated, losing strength because of corrosion and under sizing, thus they are unable to carry discharge at required pressure heads. Furthermore, increasing demands continually place additional burdens on these pipes. These problems directly contribute to the replacement/rehabilitation costs, elevated storage tank costs, and increased pumping costs. Because of the high costs involved, there is a need to identify and prioritize the failure-prone pipes for proper remedial action.

A key factor in replacing water mains is to understand the condition of the pipes in the system. Although on-site, real-time inspection of a pipe is the ideal method to analyze the condition of the pipes, this approach is expensive and cannot be practically applied to the entire distribution system. Therefore, there is a need for methods that can utilize the data collected during repairs and recommend optimal times of replacement for failure-prone pipes. A related issue is the optimal design of an expanding network with proper pipe size selection. The pipe network optimal design problem involves multiple local minima and there have been very few attempts to exploit the underlying structure of the problem to reach the global optimum. There is a need for global optimization techniques that expose the true nature of the pipe design problem. The objectives of the study are as follows:

1. Develop a failure assessment model for identifying failure-prone pipes for replacement or rehabilitation.
2. Develop a holistic global optimization model that accounts for network expansion and pipe sizing.
3. Provide an example application of the developed models

Following the above objectives, this report contains an economically sustainable critical threshold break rate derived for the replacement of failure-prone pipes. Also, a new global optimization scheme capable of solving the problem within a specified tolerance from global optimality is derived. The report is organized as follows. Chapter 2 provides information about identifying failure-prone pipes and provides a literature survey of the methods used for estimating the optimal time of replacement. In chapter 3, the economically sustainable threshold break-rate criterion is derived and its application to the Walski-Pellicia's critical break rate is shown. The break-rate criterion is also applied to the widely used Shamir-Howard's break occurrence rate to establish the optimal replacement time expression under milder conditions. The section on time-truncated pipe break-rate prediction provides the analytical Weibull ROCOF (rate of occurrence of failure) model based on truncated data. Chapter 4 contains example applications illustrating the methodology and drawing inferences related to the role of repair-replacement cost ratio, pipe diameter, and discount rate. The key findings are reported in a

summary section. Chapter 5 presents the network optimization model and derives the linear programming lower-bounding problem and the upper-bounding heuristic. A branch-and-bound algorithm to solve the problem to a specified tolerance from optimality is described in a separate section of this chapter. Chapter 6 presents extensive computational results on test problems from the literature, as well as for the Blacksburg network that is offered as another test case to researchers. Finally, Chapter 7 concludes the report with a summary of the main contributions.

CHAPTER 2

IDENTIFYING FAILURE-PRONE PIPES

2.1 Introduction

In this chapter, the available methods to identify failure-prone pipes are grouped by the underlying characteristics. It is generally believed that large diameter pipes have lower failure rates in comparison to small diameter pipes, there is a seasonal influence, and pipes fail more often at later stages. It is also understood that load multipliers and strength loss due to THE reduction in the pipes' wall thickness (attributable to corrosion) are the primary causes of failures. The underlying answer that is being sought is the optimal time of replacement (whether to continue to repair or replace). The following literature review groups methods by categories that include factors such as customer dissatisfaction, criticality or consequence, corrosion, structural weakness, load increases, the tendency for an increased break rate, and the least total cost.

2.2 Identifying Failure-prone pipes

The available methods for pipe replacement analysis can be put in the following categories: (1) Deterioration Point Assignment (DPA) schemes, (2) Breakeven Analysis, (3) Regression and Failure Probability Methods, and (4) Mechanistic Strength Assessment Methods.

2.2.1 Deterioration Point Assignment Schemes

In the deterioration point assignment schemes (DPA), a set of factors involved in pipe failure is identified. These may include the age of the pipe, pipe material, pipe size, type of soil, location, water pressure, susceptibility to frost, discoloration and odor problems, and the history of previous breaks. Ordinal descriptions of these factors are associated with numerical failure scores. For any pipe, a total failure score is obtained by adding the failure scores of the factors. If the total failure score exceeds a threshold value, the pipe is considered a candidate for replacement/repair (personal communication, Weston, Inc., 1997). The discriminatory power of the scheme is clearly limited and becomes an issue if there are other pipes competing for limited funding. Also, it is a here and now assessment and lacks the predictive power that is crucial for developing a future course of action.

2.2.2 Breakeven Analysis

The breakeven analysis is a cost-based method. It requires depositing a certain sum at an interest rate and its compounded value should be equal to the future repair and replacement costs. This appealing procedure obviously requires another predictive model for the anticipated break times to determine the needed break costs. The following is a review of such predictive schemes
.

2.2.3 Regression and Failure Probability Models

The regression and failure probability methods are related to the DPA scheme because they build on the same deterioration factors but can bring in a predictive capability by assessing the probability of survival. Comprehensive reviews of failure-related factors are given in O'Day *et al.* (1986) and Mays (2000).

Shamir and Howard (1979) applied regression analysis to obtain a relationship for the breakage rate of a pipe as a function of time. This relationship was used to find the optimal timing of pipe replacement to minimize the total cost of repair and replacement. Walski and Pelliccia (1982) provided the idea of the threshold break rate. They adopted Shamir and Howard's (1979) model for predicting break rates and derived an optimal replacement time estimator by setting the total repair costs over a period of time equal to the replacement cost.

Male et al. (1990) described a procedure in which an arbitrary threshold break rate is fixed. The analysis involved the consideration of five alternatives: (1) replace after one or more breaks, (2) replace after two or more breaks, (3) replace after three or more breaks, (4) replace after four or more breaks, and (5) the do nothing approach. Alternative 2 turned out to be the most aggressive policy. Male et al. also indicated that the choice of the alternative is sensitive to the discount rate used in the calculation. A higher rate leads to a less aggressive policy and vise versa. Male et al. (1990) drew their conclusions from simulation runs. The present work yields a closed form analytical model that also illustrates the role of the various factors identified in the paper by Male et al. (1990).

Clark et al. (1982) suggested a model that combines two equations, one to predict the time to the first break and the second to predict the number of subsequent breaks that were assumed to grow exponentially over time in an attempt to account for the relative impacts of various external agents. Clark et al. (1982) have made the following observations: only a subset of pipes has recurrent repairs; the time to the first repair is quite long, typically about fifteen years; the time between repairs becomes shorter as pipes get older; the larger diameter pipes tend to have fewer problems; and industrial development in general results in more repairs.

Kettler and Goulter (1985) provided regression equations for the number of breaks versus diameter and time for cast iron and for asbestos-concrete mains in Winnipeg, Canada. Their estimates showed a strong inverse correlation between failures and the diameter (0.0625 less annual failures/km of the main with each cm of pipe diameter, for diameters between 10 and 30 cm). The regression results for New York, Philadelphia and St. Catherines, Canada also showed decreasing failure rates with an increasing diameter in these three cities in which failures were found to increase linearly with time.

Mavin (1996) provided a review of the failure models in the literature and pointed out the need to filter the data before constructing a failure model. It was suggested not to include breaks that occurred within three years of installation and six months from a previous break repair. Based on the filtered data, a set of regression equations was constructed for the number of failures over a time period and time interval between breaks.

Marks et al. (1985) used multiple regression techniques to establish that the variables affecting the pipe breakage rate were the pipe diameter, length of the pipe section, age, pressure, type, soil corrosivity, intensity of land development, the number of previous breaks, the time to the second break, and the period of installation.

Andreou et al. (1987) applied the proportional hazards model to predict failure probabilities of pipes in the early stages of deterioration and a Poisson model for the later stages of pipe deterioration. The basic idea of this model is to estimate a survivor function for each individual pipe that will provide the probability for that pipe's survival beyond a future time period given a set of risk factors. The model provides the hazard function as a product of the baseline hazard function dependent only on time and a scaling factor made up of external variables such as pipe diameter, length, soil type, and land use.

Deb et al., (1997) discussed a probabilistic model called KANEW to estimate the miles of pipes to be replaced on an annual basis. The model uses the water main inventory, with the pipes categorized according to their age, material, diameter, and bedding quality. For each category, 100th, 50th and 25th percentile ages are obtained either by expert opinion or by an analysis. These percentiles are utilized to obtain the three parameters of the Herz probability density function from which the survival probabilities are obtained. These survival probabilities are used to obtain the expected survivors or its complement of non-survivors per year, which are to be renewed. The procedure is applied to a bundle of pipes with similar makeup as opposed to an individual pipe.

2.2.4 Mechanistic strength assessment methods

In addition to the three methods described so far, a number of researchers have developed mechanistic methods to model the pipe failure phenomena. The mechanistic methods account for the normal overburden and internal loads. They also consider the other stress multiplying factors including corrosion, temperature induced stresses, and frost load. The predictive capability is brought in either by a correlation analysis or through a probabilistic analysis by considering the parameters/variables to be random. Roberge (2000) and Agbenowosi (2000) contain comprehensive reviews.

2.3 An Assessment

The above review indicates that deterioration point assessment schemes lack a built-in mechanism to predict the future break times; also, the methods do not have an optimal procedure for replacement. The cost-based schemes provide an overall framework for the optimal replacement but must be augmented by the regression and failure probability models to project future failure times. So far, the failure models that are available in the literature are primarily based on regression and have, in general, a low correlation coefficient to predict failure times. The hazard function, strictly speaking, applies only to non-repairable systems and pipes are repairable systems. The mechanistic models address the failure causing factors explicitly to assess the condition of a pipe. The data needed to apply these models in assessing the residual strength of a pipe are hard to obtain because they pertain to the condition of the pipe underground. Agbenowosi (2000) provides a comprehensive analysis of the mechanistic models.

In this report, techniques that overcome the drawbacks of the earlier methods are presented. First, an analytical model for an economically sustainable (critical) break rate is derived. Second, the equivalence relationships between the critical break rate and statistical models suitable for repairable systems are derived. Third, the utility of the proposed procedure is established with the aid of the equivalence relationship and a flexible probability function. The methodology uses a time truncated probability function and, therefore, does not require full failure history of a pipe and accommodates an incomplete data set.

CHAPTER 3

ECONOMICALLY SUSTAINABLE THRESHOLD BREAK RATE

3.1 Introduction

In this chapter, a new methodology for optimal pipeline replacement is presented. An economically sustainable threshold break rate for the replacement of pipelines in deteriorating water distribution systems is derived. It yields some of the previously available replacement criteria under weaker restrictions. Relations of equivalence are established between the threshold break rate and the rate of occurrence of failure (ROCOF) and the hazard rate functions. These statistical functions are utilized to predict break rates for a system. Optimal replacement times are obtained by equating the threshold break rate and the assessed (predicted) break rate from the ROCOF, and the hazard rate functions. Design charts to determine the optimal threshold break rate as a function of the pipe diameter and discount rate are also given.

3.2 Threshold Break Rate Equation

At the time of the nth break, a decision has to be made whether to replace the pipe at a cost of F_n or to repair it at a cost of C_n. The scenario also implies that for the previous (n - 1) breaks only repairs have been performed. If we assume that the pipe will be replaced (C_n included in the sum should be observed to adjust F_n when necessary) at the time of nth break, t_n, we can write the present worth of the total cost of the pipe as

$$T_n = \sum_{i=1}^{n} \frac{C_i}{(1+R)^{t_i}} + \frac{F_n}{(1+R)^{t_n}} \tag{3.1}$$

in which: R = discount rate, t_i = time of ith break measured from the installation year (year), C_i = repair cost of ith break, F_n = replacement cost at time, t_n, T_n = total cost at time ‘0’ (present worth).

When a pipe is new, it experiences very few breaks. An old pipe experiences more breaks under the same trench and load conditions. Therefore, the combination of varying thetime interval between breaks (accelerated breaks towards the end), relatively smaller repair cost, and a generally large replacement cost leads to the existence of a “U” shaped present worth of the total cost curve over time. The derivation of the threshold break rate, seeks the time of the minimum present worth total cost.

For the total cost T_n at time t_n to be a minimum, assuming a unimodal function, it must satisfy the condition

$$T_{n-1} > T_n < T_{n+1} \tag{3.2}$$

However, at a point of consideration the minimum to occur the only condition of interest is $T_{n+1} > T_n$. Therefore, consider

$$T_{n+1} = \sum_{i=1}^{n+1} \frac{C_i}{(1+R)^{t_i}} + \frac{F_{n+1}}{(1+R)^{t_{n+1}}} \tag{3.3}$$

From Eqs. (3.1) and (3.3) we obtain

$$T_{n+1} - T_n = \frac{F_{n+1}}{(1+R)^{t_{n+1}}} + \frac{C_{n+1}}{(1+R)^{t_{n+1}}} - \frac{F_n}{(1+R)^{t_n}} \tag{3.4}$$

For $T_{n+1} - T_n > 0$, we have

$$\frac{F_{n+1}}{(1+R)^{t_{n+1}}} + \frac{C_{n+1}}{(1+R)^{t_{n+1}}} - \frac{F_n}{(1+R)^{t_n}} > 0 \tag{3.5}$$

Solving for $t_{n+1} - t_n$ is

$$t_{n+1} - t_n < \frac{\ln\left(\frac{C_{n+1}}{F_n} + \frac{F_{n+1}}{F_n}\right)}{\ln(1+R)} \tag{3.6}$$

Recognizing $t_{n+1} - t_n$ is the time between n th and (n + 1)th breaks or the time interval for the occurrence of one break, we obtain the threshold break rate, $Brk_{th,1}$, as the inverse of Δt_n where $\Delta t_n = t_{n+1} - t_n$. That is, the threshold break rate is defined as

$$Brk_{th} = \text{break rate between subsequent breaks} = \frac{1}{t_{n+1} - t_n} = \frac{1}{\Delta t_n} \tag{3.7}$$

Therefore, the threshold break rate is expressed as

$$Brk_{th} > \frac{\ln(1+R)}{\ln\left(\frac{C_{n+1}}{F_n} + \frac{F_{n+1}}{F_n}\right)} \tag{3.8}$$

Now, from the observed data for any given pipe, we can derive a current break rate. Whenever the current break rate, Brk_{cur} equals or exceeds Brk_{th} the pipe should be replaced. In other words, the condition for a pipe replacement at the current time is expressed as

$$Brk_{cur} \geq Brk_{th} \tag{3.9}$$

In deriving the threshold break rate in Eq. (3.8), no differentiability is assumed and the discrete break occurrences are maintained. Further, no break rate equation is embedded. The threshold break rate has the interpretation of an economically sustainable break rate and involves only the repair and replacement costs and the discount rate. Therefore, this threshold break rate, Brk_{th}, simply asserts what is not an acceptable break rate. Male et al. (1990) in their detailed simulation study drew attention to (1) the role of the higher discount rate in allowing more breaks, and (2) the somewhat invariant break costs but increasing the replacement costs for large diameter pipes favoring replacement of smaller pipes. These characteristics are readily captured in the threshold break rate given in Eq. (3.8). Walks and Felicia (1982) espoused the idea of a critical break rate and obtained it using the form of the Shamir and Howard's (1979) break rate model. They set the discounted total break cost over a number of years called the analysis period, "m", written in terms of the break rate to exceed the replacement cost. For a chosen "m", the analysis results in a break rate that will cause the cumulative discounted repair costs to exceed the replacement cost. This break rate is called the critical break rate. In this study the total present worth cost is minimized to obtain the threshold break rate. In Eq. (3.8), if we assume $F_n \approx F_{n+1}$ at an accelerated quick breaking stage and by the property of logarithm for small values, we obtain

$\ln[1 + C_{n+1}/F_n] = C_{n+1}/F_n$, which yields the special case critical break rate, $J^* = \ln(1 + R)\, F_n/C_{n+1}$ given by Walski and Pelliccia (1982).

Here, the importance of the terms 'at current time' should be emphasized. If the current break rate is less than the threshold break rate and there is a need to know when the pipe needs to be replaced, then a future break rate of the pipe must be predicted and compare the future break rate with the threshold break rate. The question arises about what sort of statistical model should be employed? In Appendix A, the equivalence relationship between the appropriate statistical functions for predicting future break rates and the threshold break rate is given.

3.3 Statistical Reliability Models

For an in depth discussion of the reliability models, refer to Ascher and Feingold (1984). The widely used statistical models for reliability analysis are the hazard function model and the rate of occurrence of failure (ROCOF) model. The hazard function is the relative rate of failure given that a component has survived up to a certain time, t, i.e., $h(t) = f_T(t)/P(T>t)$ in which $f_T(t)$ is the probability density function of the failure time T and $P(T>t)$ is the probability that the random failure time T exceeds the numerical value t. It gives the conditional probability that a component will instantaneously fail at t given that it has survived until time t, i.e., $h(t)\, dt = P(t < T \le t + dt)$. The definition of this function makes it applicable to only *non-repairable systems* with the first and only failure. Ascher and Feingold (1984) show why h(t) should not be interpreted as a conditional probability density function. The ROCOF function is defined as the rate of change of the expected number of failures given by $g(t) = dE[N_c(t)]/dt$. The ROCOF function readily applies to a *repairable system*. It measures the rate at which the number of failures is changing over time. Water mains are clearly repairable systems. The data also readily permits one to evaluate the ROCOF function as shown in the examples in Chapter 4. Appendix A shows the equivalence between the ROCOF and hazard functions and the threshold break rate of Eqs.(3.7) and (3.8). In the ensuing section the threshold break rate is set equal to the Shamir and Howard's exponential break rate to obtain the optimal replacement time. This end result is derived with less restrictive assumptions.

3.4 Application to the Shamir and Howard's Exponential Break Rate Model

Consider the break rate equation (Shamir and Howard, 1979)

$$N(t) = N(t_0)e^{A(t-t_0)} \tag{3.10}$$

where: N(t) = number of breaks per 1000 ft length of pipe in year t, t = time in years, t_0 = base year for the analysis(pipe installation year, or the first year for which data are available), A = growth rate coefficient (1/year). In our notation, N(t) is the break rate (that is, number of breaks/year at year t). Therefore, based on Appendix A, we set $N(t) = Brk_{th}$ given in Eq. (3.8)

$$N(t) = N(t_0)e^{A(t-t_0)} = \frac{\ln(1+R)}{\ln\left(\frac{C_{n+1}}{F_n} + \frac{F_{n+1}}{F_n}\right)} \tag{3.11}$$

Assuming $F_{n+1} \approx F_n$ at a rapidly deteriorating stage in which breaks occur in quick succession, we have

$$N(t_0)e^{A(t-t_0)} = \frac{\ln(1+R)}{\ln\left(1+\frac{C_{n+1}}{F_n}\right)} \tag{3.12}$$

Further for small x, putting

$$\ln(1+x) = x \tag{3.13}$$

we obtain

$$N(t_0)e^{A(t-t_0)} = \frac{\ln(1+R)F_n}{C_{n+1}} \tag{3.14}$$

from which

$$t^* = t_0 + \frac{1}{A}\ln\left[\frac{\ln(1+R)F_n}{N(t_0)C_{n+1}}\right] \tag{3.15}$$

which is the same as Shamir and Howard's end result but obtained without any differentiation and maintaining the discrete nature of the break events. Kleiner et al. (1998) embedded Shamir and Howard's break rate model, Eq.(3.10), in a rolling time horizon analysis to obtain an optimal time of replacement. Further, Kleiner and Rajani (1999) applied the model with the aid of regression equations. They also formalized how to proceed with a replacement analysis for water utility as a whole.

3.5 Time Truncated Pipe Break Rate Prediction with the (ROCOF) Function

In Appendix A, the equivalence between the ROCOF (Rate of occurrence of failure) function and the threshold break rate is shown. This function not only helps to project the future break rate but also the related probabilities can be computed. It is also possible to bring in the environmental factors that influence failure in the form of a regression within the probabilistic ROCOF analysis. In the following the needed theoretical structure is established. A good reference for the description given below is Bain and Engelhardt(1991). One of the properties of the homogeneous Poisson process is that the times between successive failures are independent random variables, each having an exponential density with parameter λ (constant ROCOF) so that

$$\Pr\{\text{Failure Time} > x\} = e^{-\lambda x}, \quad 0 < x < \infty. \tag{3.16}$$

In a nonhomogeneous Poisson process (NHPP) with time-dependent ROCOF, l (t), the number of failures in the time interval (t_1, t_2] has a Poisson distribution with mean

$$\int_{t_1}^{t_2} \lambda(t)dt. \tag{3.17}$$

Thus the probability of no failures in (t_1, t_2] is

$$\exp\left\{-\int_{t_1}^{t_2} \lambda(t)dt\right\}. \tag{3.18}$$

By choosing a suitable parametric form for $\lambda(t)$, a flexible model for the failures of a repairable system is obtained.

If a system is observed for the time interval $(0, t_0]$ with given n failures occurring at times $t_1, t_2, \ldots, t_n$, the zero coincides with the beginning of the observation time for a pipe and thus accommodates truncated data (see Bain and Engelhardt, 1991). The joint probability of observing no failures in $(0, t_1)$, one failure in $(t_1, t_1 + \Delta t_1)$, no failures in $(t_1 + \Delta t_1, t_2)$, one failure in $(t_2, t_2 + \Delta t_2)$ and so on up to no failures in $(t_n + \Delta t_n, t_0)$ is given by (for small $\Delta t_1, \Delta t_2, \ldots, \Delta t_n$)

$$\left\{\exp\left(-\int_0^{t_1}\lambda(t)dt\right)\right\}\lambda(t_1)\Delta t_1\left\{\exp\left(-\int_{t_1+\Delta t_1}^{t_2}\lambda(t)dt\right)\right\}\lambda(t_2)\Delta t_2\ldots\left\{\exp\left(-\int_{t_n+\Delta t_n}^{t_0}\lambda(t)dt\right)\right\}. \tag{3.19}$$

Dividing through by $\Delta t_1 \Delta t_2 \ldots \Delta t_n$ and letting $\Delta t_i \to 0$ (i = 1, 2, …, n) gives the likelihood function.

$$L = \left\{\prod_{i=1}^{n}\lambda(t_i)\right\}\left\{\exp\left(-\int_0^{t_0}\lambda(t)dt\right)\right\}. \tag{3.20}$$

In this study, the Weibull form of the ROCOF function is used because of its wide use in repairable systems modeling. Weibull's ROCOF function can be expressed as

$$\lambda(t) = \gamma\delta t^{\delta-1}. \tag{3.21}$$

Therefore, for Weibull's ROCOF function, the maximum likelihood estimates for g and d are obtained as

$$\hat{\delta} = \frac{n}{n\log t_0 - \sum_{i=1}^{n}\log t_i}. \tag{3.22}$$

and

$$\hat{\gamma} = \frac{n}{t_0^{\hat{\delta}}}. \tag{3.23}$$

Once the parameters of the model are determined, the threshold break rate equation can be used to obtain the optimal replacement time for each pipeline. Thus,

$$\gamma\delta(t)^{\delta-1} = \frac{\ln(1+R)}{\ln\left(1+\frac{C}{F*L}\right)}. \tag{3.24}$$

where: F = cost per unit length, and L = replacement pipe length. Therefore, optimal replacement time is obtained as

$$t^* = \left(\frac{\ln(1+R)}{\gamma\delta\ln\left(1+\frac{C}{F*L}\right)}\right)^{\frac{1}{\delta-1}}. \tag{3.25}$$

3.6 Summary

In this chapter a threshold break rate is derived. In contrast to the previous studies the derivation does not embed a rate of break occurrence model. The threshold break rate reflects the important conclusions drawn in terms of the discount rate and repair to the replacement cost ratio by Male et al. (1990). It yields an optimal time of replacement as obtained by Shamir and Howard (1979) but with less restrictive assumptions. It also reproduces a special critical break rate given by Walski and Pelliccia (1982). Further more, the equivalence between the threshold break rate and the statistical failure modeling functions of rate of occurrence of failure and hazard functions is established. These functions provide a rich modeling environment for predicting the pipe break rate from a break database. By setting the threshold break rate to be equal to the projected pipe break rates from the ROCOF and the hazard rate analyses, optimal replacement time expressions are obtained.

3.7 Appendix A: Equivalence relationship between the threshold break rate and statistical failure rate models

The relationship between the Threshold Break Rate, the Rate of Occurrence of Failure (ROCOF) Function and the Hazard Function

A break rate function is defined as (Ascher and Feingold, 1984)

$$r(t) = \lim_{\Delta t \to 0} \frac{\text{No. of breaks in } (t, t + \Delta t]}{\Delta t} \qquad \text{(A.1)}$$

Equation (A.1) is also expressed as the derivative of the expected cumulative number of breaks function $E[N_c(t)]$, that is

$$r(t) = \frac{dE[N_c(t)]}{dt} \qquad \text{(A.2)}$$

In general, Eq. (A.2) is called the rate of occurrence of failure (ROCOF) function. From Eq. (A.1) an empirical break rate in $(t_n, t_n + \Delta t_n]$, where t_n is the nth break time, and $\Delta t_n = t_{n+1} - t_n$, can be obtained as

$$r(t_n) = \frac{N_c(t_{n+1}) - N_c(t_n)}{\Delta t_n} \qquad \text{(A.3)}$$

Because t_{n+1} and t_n are successive failure times, we have $N_c(t_{n+1}) - N_c(t_n) = 1$ and the empirical break rate becomes

$$r(t_n) = \frac{1}{\Delta t_n} \qquad \text{(A.4)}$$

which is the same as the threshold break rate given in (7), and

$$r(t_n) = Brk_{th} \qquad \text{(A.5)}$$

The hazard function called the hazard rate or the probability of instantaneous failure rate is defined by

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr[t < T \le t + \Delta t \mid T > t]}{\Delta t} \tag{A.6}$$

where T is the failure time random variable. The hazard function expresses the propensity to fail in the next small interval of time, given survival to time t. That is, for small Δt,

$$h(t) \cdot \Delta t \approx \Pr[t < T \le t + \Delta t \mid T > t] \tag{A.7}$$

The hazard function originally applies to non-repairable systems in which a failure implies the death of a system and is allowed only once in its lifetime. However, to apply the hazard function, we assume here that a system gains a new life after each repair. Similar to the case of the break rate function, the estimate of the hazard function at time t_n is expressed as

$$h_e(t_n) = \frac{N_c(t_{n+1}) - N_c(t_n)}{N_c(t_n)\Delta t_n} \tag{A.8}$$

where the numerator is interpreted as the number of deaths in Δt_n and $N_c(t_n)$ is the number of survivors at time t_n. Now, consider the situation in which we are continuously monitoring a pipe for every break. In such a case

$$N_c(t_{n+1}) - N_c(t_n) = 1 \tag{A.9}$$

and for $N_c(t_n) = 1$, the estimate of the hazard function at time t_n is

$$h_e(t_n) = \frac{N_c(t_{n+1}) - N_c(t_n)}{N_c(t_n)\Delta t_n} = \frac{1}{\Delta t_n} \tag{A.10}$$

which is the hazard rate of the lone survivor or perhaps, the extinction rate. Eq. (A.10) has the same definition as the threshold break rate shown in Eq. (7). Therefore, the optimal replacement time of a pipe can be obtained by solving the equation

$$h(t) = Brk_{th} \tag{A.11}$$

for time, t. The equivalence relationships established in this Appendix enable the reader to take advantage of the rich development in the area of life testing and statistical failure modeling.

CHAPTER 4

APPLICATION OF THE THRESHOLD BREAK-RATE

4.1 Examples

In this chapter three examples that use the same break time data are given. Example 1 relates the threshold break rate to the *empirical* ROCOF function to obtain the optimal replacement time. In Example 2, the hazard function, $\lambda(t)$, is considered. In Example 3, a Weibull ROCOF function is fitted.

Example 4.1: Given the average times of failure for the nth break in Table 4.1, an annual discount rate of 7.5%, a repair cost of $3,000 and a replacement cost of $100,000, compute the optimal replacement time.

Solution: The break times satisfy the equation $T(n) = 50(1-0.8^n)$. From this equation we obtain the break rate as $dn/dt = 0.08963/0.8^n$ which depends on the break number. From our analysis we know that the break rate

$$dn/dt = 0.08963/0.8^n = Brk_{th} = \frac{\ln(1+R)}{\ln\left(\frac{C_{n+1}}{F_n} + \frac{F_{n+1}}{F_n}\right)} = 1.844$$

which gives the optimal break number as $n^* = 14$. Therefore, the replacement time is 47.8 years.

Example 4.2. Given the times of failure for the nth break in Table 4.1, an annual discount rate of 7.5%, a repair cost of $3,000, and a replacement cost of $100,000, compute the optimal replacement time in terms of the hazard function.

Solution: Table 4.1 shows the times between breaks obtained from the successive differences of the break times. Assuming the inter-break times follow exponential distributions, we obtain the exponential probability density parameter given in column 4 which is the reciprocal of the expected inter-break time. It is also true that the hazard function for exponential probability distribution is its parameter *Alpha*. As it is shown in Appendix A, the hazard function of a lone survivor equals the threshold break rate [see Eq. A(10)] we obtain $Brk_{th} = \frac{\ln(1+R)}{\ln\left(\frac{C_{n+1}}{F_n} + \frac{F_{n+1}}{F_n}\right)} =$ 1.844. The optimal replacement time from Table 4.1 is around 47.8 which belongs to the hazard 1.819 at the break number 14.

Example 4.3: Given the times of failure for the nth break in Table 4.2 (subset of Table 4.1), an annual discount rate of 7.5%, a repair cost of $3,000, and a replacement cost of $100,000, compute the optimal replacement time using the *Weibull* ROCOF function.
Solution:

First of all, in Table 4.2, the break times are considered after the occurrence of the eighth break at 41.6114 to obtain the best fit. It also illustrates that it is not necessary to have the complete data for the analysis and we can use the truncated data. The majority of the water utilities does not have complete data sets. The second column in Table 4.2 gives the break times since the eighth break. Using the break times after the eighth break, we have d = 2.212 from Eq. (3.22) and g = 0.1066 from Eq.(3.23) for n = 7 and t_0 = 6.6294. Therefore, the optimal replacement time from Eq. (3.25) is 5.455 years since the 8th break or 41.6114 + 5.455 = 47.076 years from the beginning which is close to the results of Examples 1 and 2. Figure 4.1 shows the plot of the fitted Weibull ROCOF function and its empirical estimate.

Discussion of the Examples: The purpose of the examples is twofold: (1) to illustrate the utility of the proposed break rate and (2) to verify the accuracy of the equivalence between the statistical functions and the threshold break rate. To elaborate on the second aspect, consider the examples. In example 1, the optimal replacement time is computed from dn/dt, where n(t) is the cumulative number of failures at time t. In example 2, the hazard function is computed by $1/\Delta T_n$ which is an entirely different estimator from dn/dt used in example 1. In example 3, yet another estimator as determined from the maximum likelihood method [see Eqs. (3.22) and (3.23)] using a subset of break times is employed in the Weibull ROCOF function. But for the equivalence relationship established in Appendix A, we would not expect the three different manipulations of the data to yield the same result.

For real systems, defining a suitable pipe length for observing break data can be difficult. As a practical matter, it is recommended to choose a pipe length that will provide a good fit based on the observed repairs by a trial and error process. Park (2000) suggests regression as a way to ascertain the break trend (also, see Kleiner and Rajani, 1999). For statistical concepts related to parameter estimation, refer to Bain and Engelhardt (1991). Walski (1984, chapter 7) offers broad-scale statements about general pipe replacement policies. For practical use, the methodology clearly helps to prioritize water mains for replacement. The actual replacement of ranked pipes will depend on the available budget and other considerations of an involved utility. At present many utilities are gathering data on pipe failure times. This practice will improve the usefulness of the analytical models such as the ones presented here.

4.2 Role of Cost Ratio and Pipe Diameter

The analysis presented in this section depends on the specific cost ratio, CR, relationship obtained with the cost data reported in Table 4.3. As the CR equation changes, the inferences drawn should change as well. The main intention is to draw attention to the types of observations that can be made from simple break rate versus length plots (see Figs. 3.2 and 3.3). The threshold break rate is expressed as

$$\mathrm{Brk}_{\mathrm{th}} = \frac{\ln(1+\mathrm{R})}{\ln\left(1+\dfrac{\mathrm{C}}{\mathrm{F}*\mathrm{L}}\right)} \tag{4.1}$$

where: F = replacement cost per unit length of a pipe ($/ft) and L is the length of pipe (ft). Considering that pipes are easily categorized by size, Eq. (4.1) is modified to represent the

threshold break rate for different diameter pipes. If a linear relationship is assumed between the diameter and the cost ratio (C/F) given by C/F = CR = A*D + B
where: A and B are regression coefficients and D is diameter of pipe, then Eq. (4.1) is expressed as

$$Brk_{th} = \frac{\ln(1+R)}{\ln\left(1+\frac{A*D+B}{L}\right)} \tag{4.2}$$

where: L is the length of pipe.

Table 4.3 shows replacement cost per unit length (ft) and repair cost per break incident used in this study. In this study, only 6- through 12-inch diameter pipes are considered because they constitute the bulk of the system's pipe inventory. The fitted linear equation using the cost ratios from Table 4.3 is CR = 74.056 D – 7.204, where: D is the diameter (ft) of pipe and CR is the cost ratio (repair cost/replacement cost per foot). Therefore, Eq. (4.2) is expressed as

$$Brk_{th} = \frac{\ln(1+R)}{\ln\left(1+\frac{74.056*D-7.204}{L}\right)} \tag{4.3}$$

By using Eq. (4.3), a series of graphs can be generated to determine the threshold break rate for different size pipes for given length and discount rate. Figures 3.2 and 3.3
show examples of such graphs.

From Table 4.3, it is seen that the cost ratio C/F increases with increasing pipe diameter. Therefore, mathematically as C/F increases for the same fixed discount rate, the threshold rate should decrease. According to Figure 4.2, a 10-inch pipe should be replaced when the break rate (breaks/year) reaches 5 for a given length of 4000 ft and the discount rate of 0.06. On the other hand the threshold break rate of an 8-inch pipe is shown to be about 5.5 given the same length and the discount rate.

There are two possibilities. Assume that larger diameter pipes do have smaller break rates from observed break data in comparison with the smaller diameter pipes at the corresponding times. This result implies that one should wait for a longer period for a bigger pipe to reach a certain threshold break rate than a smaller one for replacement. One should not confuse threshold break rate with optimal replacement time. On the other hand, if a larger pipe does have frequent breaks, then the cost ratio implies that it is prudent to replace it at a lower threshold break rate value in comparison to a smaller pipe.

For example, consider 10-inch and 12-inch pipes with a threshold break rate of 5. Figure 4.2 shows that while 4000 ft of 10-inch pipe should be replaced when it reaches a threshold break rate of 5, 5000 ft of 12-inch pipe should be replaced when it reaches the same break rate. This is because larger pipes fail so infrequently that longer lengths of pipes should be considered to have the same break rate as a smaller pipe. Another interpretation is that for the same given length larger diameter pipes will have to be replaced at a lower threshold rate to avoid frequent, larger repair costs. As observed earlier, an increase in the discount rate increases the threshold break

rate. Figure 4.2 shows a break rate of 2 for 1000 ft-long 6-inch pipe for a discount rate of 6% whereas Figure 4.3 yields about 2.5 breaks for a discount rate of 7% for the same pipe.
An interesting aspect of this regression analysis is the criterion to replace a 1000 ft. section for 3 breaks per year proposed by Morris (1975) as an empirical finding and reported in Walski and Pelliccia (1982) [also see Andreou et al. (1987)]. In the present analysis, the cutoff number of breaks per year varies between 2 and 3 for a six-inch pipe; and in general the cut off number of breaks per year depends on repair and replacement cost ratio, diameter, and discount rate. It must also be observed that Figs. 3.2 and 3.3 are based on the data given in this report. They are meant to serve as example charts for performing certain sensitivity analyses such as the inferences drawn above. An interested reader should create one's own charts with the applicable data.

4.3 Summary

In this chapter the utility of the threshold break rate derived in Chapter 3 is illustrated by several examples. Example calculations verify the accuracy and the equivalence relationships established in Appendix A of Chapter 3. A derived cost-ratio relationship between the repair and replacement costs helps to create practical plots between pipe length and the threshold break rate. These plots enable an engineer to decide on the failure nature of the large and small diameter pipes as well as the effect of the discount rate.

Table 4.1. Break Data

Break, n	Break Time, T_n	Inter Break Time, ΔT_n	Alpha(n)
1	10.0000	10.0000	0.1000
2	18.0000	8.0000	0.1250
3	24.4000	6.4000	0.1563
3	29.5200	5.1200	0.1953
5	33.6160	3.0960	0.2441
6	36.8928	3.2768	0.3052
7	39.5142	2.6214	0.3815
8	41.6114	2.0972	0.4768
9	43.2891	1.6777	0.5960
10	44.6313	1.3422	0.7451
11	45.7050	1.0737	0.9313
12	46.5640	0.8590	1.1642
13	47.2512	0.6872	1.4552
14	47.8010	0.5498	1.8190
15	48.2408	0.4398	2.2737
16	48.5926	0.3518	2.8422
17	48.8741	0.2815	3.5527
18	49.0993	0.2252	3.4409
19	49.2794	0.1801	5.5511

Alpha(n) = $1/\Delta T_n$ (1/year)

Table 4.2. Break Times for Weibull ROCOF

Break Number, n	Break Time	Time since the 8th break
9	43.2891	1.6777
10	44.6313	3.0199
11	45.7050	4.0936
12	46.5640	4.9526
13	47.2512	5.6398
14	47.8010	6.1896
15	48.2408	6.6294

Table 4.3. Cost Table by Pipe Size

Size(inch)	Replacement Cost ($/ft), F	Repair Cost($), C	Cost Ratio C/F
6	92.77	2814.00	30.333
8	96.95	3985.00	41.104
10	106.50	5869.00	55.108
12	116.05	7753.00	66.807

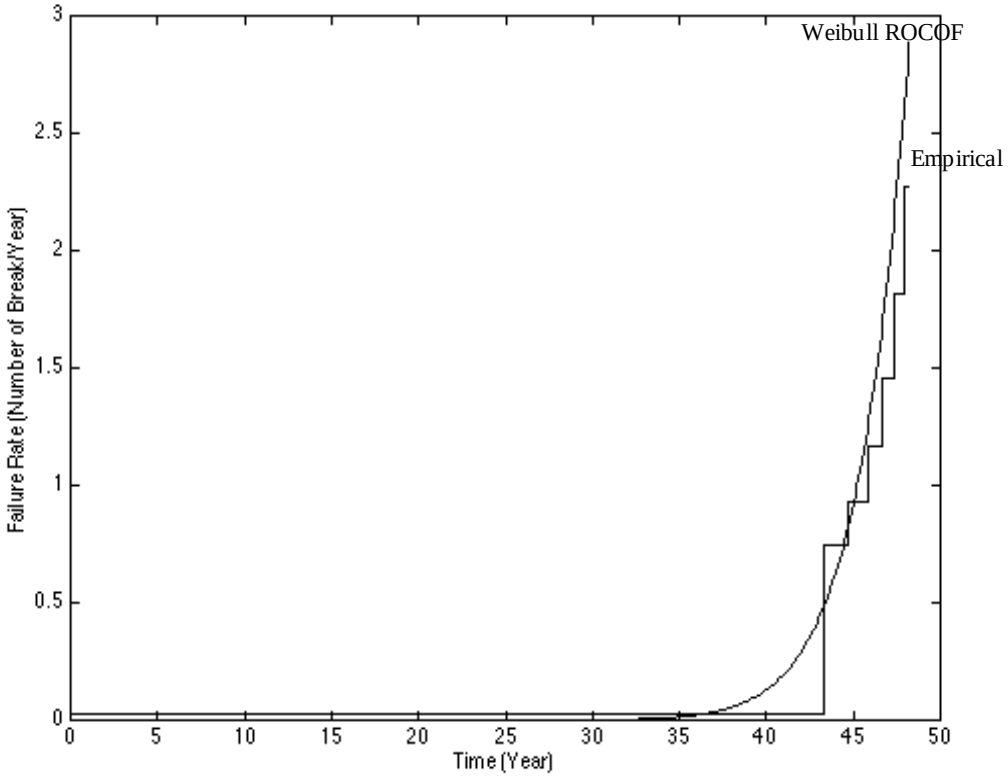

Figure 4.1. Weibull ROCOF function

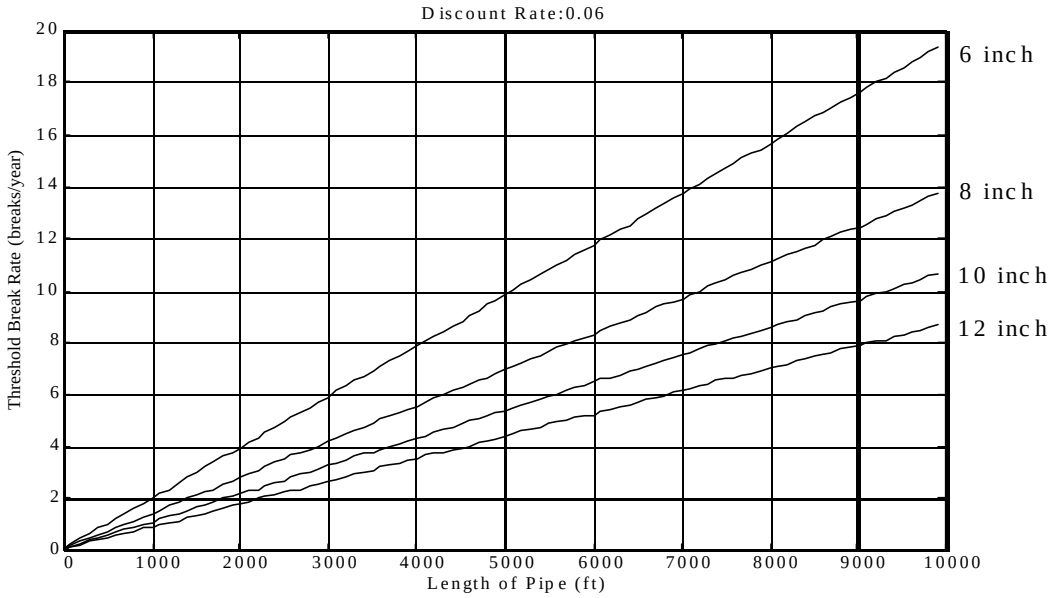

Figure 4.2. Threshold Break Rate Plot (R = 0.06, CR= 74.056 D– 7.204)

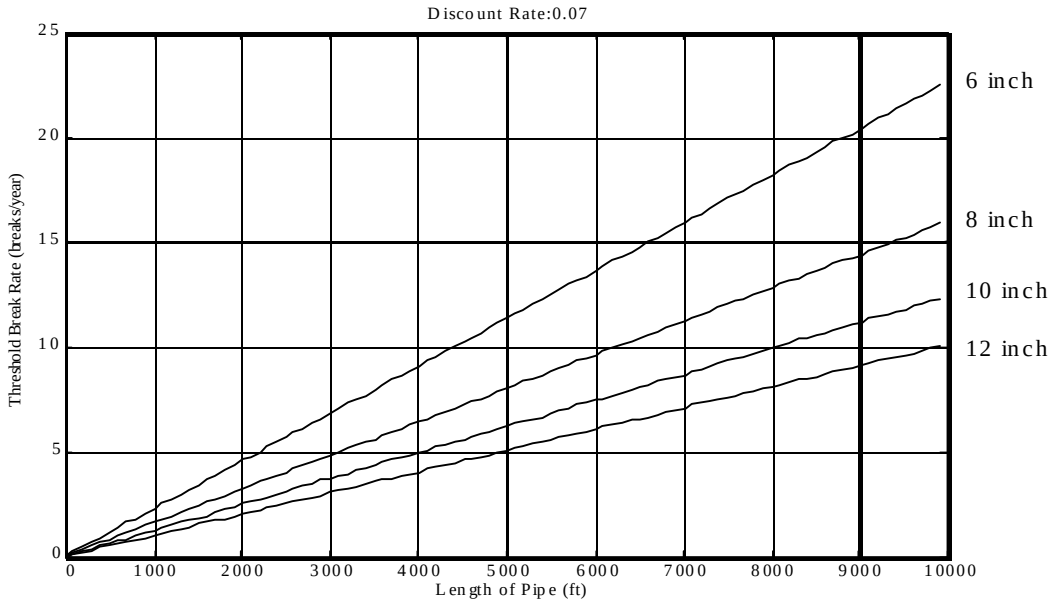

Figure 4.3. Threshold Break Rate Plot (R = 0.07, CR= 74.056 D– 7.204)

CHAPTER 5

GLOBAL OPTIMIZATION MODEL FORMULATION

5.1 Introduction

The investments associated with the installation, expansion, and maintenance of water distribution systems are very high, and account for a significant proportion in municipal maintenance budgets. An important component in this process of designing a cost effective water distribution system or extending a pre-existing network is to design the paipes of various sizes that are capable of satisfying the flow demand, in addition to satisfying the minimum pressure head and hydraulic redundancy requirements. However, this least cost pipe design problem is a hard nonconvex optimization problem having a number of local optima, and has proven difficult to solve. A number of research efforts over the last two decades have focused on solving this problem, most of them generating improved suboptimal solutions for several standard test problems from the literature, with no adequate lower bounds to evaluate the prescribed solutions. Three notable exceptions discussed below that are capable of providing solutions within a proven tolerance of a global optimum are the methods due to Eiger *et al.* (1994), Sherali and Smith (1995), and Sherali, *et al*. (1998). Our development in the present chapter is a further enhancement of these three procedures. For other related expositions, we point the reader to the survey by Lansey and Mays (1989), the holistic integrated pipe-reliability-and-cost network optimization approach and implementation discussion in Sherali and Smith (1993), and the recent application of genetic algorithms as described in Dandy *et al*. (1996).

5.2 Literature Review

A first global optimization approach to the least cost pipe sizing decision was proposed by Eiger *et al*. (1994). A branch-and-bound algorithm is developed in this chapter, based on partitioning the hyperrectangle restricting the flows into several subrectangles. At each node of the branch-and-bound tree, a subgradient-based heuristic is applied to determine an upper bound via the non-smooth, non-convex, projection of the problem onto the space of the flow variables. An independent relaxed, duality-based linear programming formulation is used to compute lower bounds.

Sherali and Smith (1995) present another global optimization approach for an arc-based formulation of the problem, in contrast with the loop-and-path based formulation employed by Eiger *et al*. (1998). They employ a Reformulation-Linearization Technique (RLT) to construct tight linear programming relaxations for the given problem in order to compute lower bounds. The procedure is embedded in a branch-and-bound scheme. Convergence to an optimal solution is induced by coordinating this process with an appropriate partitioning scheme. Several test problems from the literature are solved to exact global optimality for the first time using this approach. In particular, these results indicate that some of the solutions reported by Eiger *et al*. (1994) are in error due to a degree of infeasibility in the flow conservation constraints.

Sherali *et al*. (1998) provide an alternative global optimization approach that enhances the method of Eiger *et al*. (1994), thereby enabling them to solve the Hanoi network test problem to within 0.486% of optimality. This was a significant improvement upon the previously best

solution reported in the literature. In this procedure, they derive a linear lower-bounding problem by relaxing the nonlinear constraints in the transformed space via polyhedral outer approximations. Upper bounds are computed by solving a projected linear program which uses the flow conserving solution generated by the lower-bounding problem. These bounding strategies are embedded within a branch-and-bound algorithm. A partitioning scheme is employed that induces a convergent process toward a feasible solution that lies within any prescribed accuracy tolerance of global optimality.

The global optimization method addressed in the present chapter provides a further enhancement of Sherali *et al*.'s (1998) polyhedral outer approximation scheme by way of (a) employing tighter relaxations; (b) using a maximal spanning tree-based projected-space partitioning scheme that dramatically reduces the computational effort, and (c) a more effective branching variable selection strategy. An alternative RLT-based relaxation that was more effective than the one proposed by Sherali and Smith (1995) was also developed and tested for comparison purposes. The foregoing enhancements are shown to substantially reduce the computational effort while determining proven global optimal solutions that lie at least within 10^{-4} % of optimality to standard test problems available in the literature. The results provide improved incumbent solutions over those previously reported in the literature for all these problems, and particularly so for the two larger Hanoi and New York City problems. In fact, for the latter problem, not even a reasonable lower bound had previously been reported. Aa tight global lower bound for this problem was derived for the first time, solving this test case to within a 10^{-6} % (or $ 0.4) of optimality.

The remainder of this chapter is organized as follows. The model formulation section presents the network optimization model, and the section on lower and upper-bounding problems derives the proposed linear programming lower-bounding problem and the upper-bounding heuristic. The branch-and-bound algorithm is described in the subsequent section.

5.3 Model Formulation

Consider a distribution network $G'(N', A')$ comprised of a set of reservoirs or supply nodes and a set of consumption or demand nodes. Let these nodes be collectively identified by the index set $N' = \{1, 2, \ldots, n\}$, where the set of source nodes is denoted by $S' \subset N'$ and the set of demand nodes is denoted by $D' \subset N'$ such that $N' = S' \cup D'$. Associate with each node a quantity b_i that represents the net water supply rate or demand rate corresponding to node *i* in the index set N'. Assume that $b_i > 0$ for $i \in S'$ and $b_i \leq 0$ for $i \in D'$. To ensure feasibility, we assume that the total supply rate is at least equal to the total demand rate.

For each pipe (new or existing) that connects a certain designated node pair *i* and *j*, where *i*, $j \in N'$, $i < j$, a (notationally) directed arc $(i, j) \in A'$ is created. For each $(i, j) \in A'$, let L_{ij} denote the pipe length corresponding to a connection in the network between the nodes *i* and *j*. If we are working on the more general problem of expanding an existing network, the problem becomes one of designing new connections, as well as, constructing parallel pipe links that need to be installed between certain specified node pairs *i* and *j*, similar to the consideration of Loganathan *et al*. (1995). Let $A_f \subseteq A'$ denote existing pipes that will remain fixed in design, and let $A_r \subseteq A'$ denote existing pipes for which a parallel replacement connection needs to be

designed. The flow between the node pairs associated with the links in A_r will then be carried by the previously existing pipe link, as well as by the newly designed parallel pipe link. (When the existing pipe is being *replaced*, we simply treat this case similar to that of designing a new connection between the corresponding nodes, and accordingly absorb this within the arc set $A' - (A_f \cup A_r)$.) For the pth such node pair $(i, j) \in A_r$, $p = 1, ..., |A_r|$ in order to avoid parallel multi-arcs, we represent the existing arc as two contiguous arcs by creating a dummy node $n + p$, along with arcs $(i, n + p)$ and $(n + p, j)$. We set the demand value b_{n+p} for the dummy node to be zero. The lengths of the arcs $(i, n + p)$ and $(n + p, j)$ are each set equal to $L_{ij}/2$, and the diameters of the segments in these pipes are fixed at their existing values. It is assumed that the "installation" cost for such pre-existing pipes is zero. Under this scheme, the corresponding link to be newly designed is now represented by the arc (i, j) that runs parallel to the arcs $(i, n + p)$ and $(n + p, j)$. Such a procedure adds $|A_r|$ nodes and $2|A_r|$ arcs to the network G', effectively creating an expanded network $G(N, A)$, where $N = N' \cup \{n + p : p = 1, ..., |A_r|\}$, $A = A' \cup \{(i, n + p), (n + p, j) : (i, j) \in A_r\}$. The revised set of demand nodes D associated with this network G are suitably updated to include the newly created (zero-demand) nodes, so that $D = D' \cup \{n + p : p = 1, ..., |A_r|\}$. Since no new source nodes are added, we have $S = S'$. For notational convenience, we will denote the set of arcs in A that are to be *newly designed* as P, while $A - P$ will represent the *existing links* in the network. We will assume that each link that needs to be designed is constructed from segments of lengths having standard available diameters, chosen from the set $\{d_k, k = 1, ..., K\}$. Also, let us denote by c_k the cost per unit length for a pipe of diameter d_k.

Associated with each link connecting node pairs $(i, j) \in A$ is the decision variable q_{ij} that represents the flow rate (m^3/hr). Note that this variable may be nonnegative or negative, thus permitting flow in either direction. A positive flow value means that flow is along the specified conventional direction of the arc. The value q_{ij} associated with each link is assumed to lie between some analytically determined minimum and maximum bounds $q\min_{ij}$ and $q\max_{ij}$, that may appropriately be of either sign. We define the hyperrectangle restricting the flows $\boldsymbol{q}$ as $W = \{\boldsymbol{q} : \boldsymbol{q}\mathbf{min} \le \boldsymbol{q} \le \boldsymbol{q}\mathbf{max}\}$, where the notation $\boldsymbol{q}\mathbf{min}$ and $\boldsymbol{q}\mathbf{max}$ with the subscripts dropped denotes the corresponding vectors of lower and upper bounds. Sherali *et al*. (1998) discuss procedures for determining these bounds on the flows from the network configuration using logical arguments (without making any *a priori* assumption on the nature of the optimal flow distribution).

The next set of decision variables relates to the lengths of pipe segments having different standard diameters, that comprise each link of the network. Let x_{ijk} denote the length of segment of diameter k in the link $(i, j) \in A$, and let x_{ij} be the vector having components $(x_{ijk}, k = 1,\ldots,K)$. We assume that the variable x_{ijk} is fixed at a value $\tilde{x}_{ijk}$ for all arcs $(i, j) \in A - P$, $\forall k = 1, ..., K$.

Now consider the energy heads at the various nodes in the network. For each node $i \in N$, let E_i denote its ground elevation, and let H_i (a variable) denote the established head above E_i. Additionally, for the source nodes $i \in S$, let F_i denote the fixed maximum available energy

head, and suppose that there is an opportunity to further raise this head by an amount H_{si} at an annualized cost $c_{si} > 0$ per unit energy head, as suggested by Rowell and Barnes (1982). Correspondingly, for each demand node $i \in D$, suppose that there is the requirement that at a flow equilibrium, the established head $(H_i + E_i)$ at this node lies in the interval $[H_{iL}, H_{iU}]$ where $H_{iL} < H_{iU}$. For any dummy node *n+p* that is formed by the conjunction of the arcs $(i, n + p)$ and $(n + p, j)$ $\forall (i, j) \in A_r$, the node elevation E_{n+p} is set equal to $(E_i + E_j)/2$, and the pressure bounding interval for $(H_{n+p} + E_{n+p})$ is set equal to $[\min\{H_{iL}, H_{jL}\}, \max\{H_{iU}, H_{jU}\}]$.

The pressure loss (or head loss) in a pipe due to friction, given by $[(H_i + E_i) - (H_j + E_j)]$ for a link (i, j), can be described by the empirical Hazen-William equation as follows (see Walski 1984), where the sign depends on the direction of flow.

$$F(q, C_{HW}, d, x) = (1.52)\ 10^4\ \text{sign}(q)\ |\ q / C_{HW}\ |^{1.852}\ d^{-4.87} x \tag{5.1}$$

where
F = pressure head loss (in meters) assuming smooth flow conditions in a given pipe segment,
q = water flow rate in the pipe (m^3/hr),
C_{HW} = Hazen-Williams coefficient based on roughness and diameter,
d = pipe diameter (in centimeters),
x = pipe length (in meters).

For our model, the head loss in a pipe that has several potential segments of varying diameter and roughness is computed as follows

$$F_{ij}(q_{ij}, \mathbf{x}_{ij}) = \sum_{k=1}^{K} F(q_{ij}, C_{HW(ijk)}, d_k, x_{ijk}), \text{ where } \mathbf{x}_{ij} \equiv (x_{ijk}, k = 1, ..., K). \tag{5.2}$$

The network optimization problem NOP, restricted on *W*, can now be formulated as follows.

NOP(*W*):

$$\text{Minimize} \quad \sum_{(i,j)\in P} \sum_{k=1}^{K} c_k x_{ijk} + \sum_{i \in S} c_{si} H_{si} \tag{5.3a}$$

subject to

$$F_{ij}(q_{ij}, \mathbf{x}_{ij}) = (H_i + E_i) - (H_j + E_j) \qquad \forall\ (i, j) \in A \tag{5.3b}$$

$$\sum_{k=1}^{K} x_{ijk} = L_{ij} \qquad \forall\ (i, j) \in P \tag{5.3c}$$

$$\sum_{j \in FS(i)} q_{ij} - \sum_{j \in RS(i)} q_{ji} = b_i \qquad \forall\ i \in D \tag{5.3d}$$

$$\sum_{j \in FS(i)} q_{ij} - \sum_{j \in RS(i)} q_{ji} \le b_i \qquad \forall\ i \in S \tag{5.3e}$$

$$q\min_{ij} \le q_{ij} \le q\max_{ij} \qquad \forall\ (i, j) \in A \tag{5.3f}$$

$$H_i + E_i \le F_i + H_{si} \qquad \forall\ i \in S \tag{5.3g}$$

$$H_{iL} \le H_i + E_i \le H_{iU} \qquad \forall\ i \in D \tag{5.3h}$$

$$H_{si} \ge 0 \qquad \forall\ i \in S \tag{5.3i}$$

$$x_{ijk} \ge 0 \qquad \forall\ (i, j) \in P,\ k = 1, ..., K$$

$$x_{ijk} = \tilde{x}_{ijk} \qquad \forall\ (i, j) \in A - P,\ k = 1, ..., K. \tag{5.3j}$$

The objective function, Equation (5.3a), denotes the total cost of the pipes and the cost of the additional head generated at each source node. Constraints (5.3b) are the conservation of energy equations, and along with Constraints (5.3h), ensure that the hydraulic energy loss over each chain in the network is such that the minimum head requirements (H_{iL}) are met for each demand node. It may be noted that Constraints (5.3b) implicitly enforce that the hydraulic energy loss in each loop in the network is zero. The link length constraints are represented by Equation (5.3c). Equations (5.3d) and (5.3e) enforce the conservation of the flow at all nodes, and Equation (5.3f) bounds the flow value in each link to lie in a specified valid or implied interval. These bounds that define W will be suitably modified during the course of the algorithm for solving Problem NOP. Constraints (5.3g) and (5.3h) represent restrictions on the maximum variable head at each source node, and the head requirements at each demand node, respectively. Finally, Constraints (5.3i) and (5.3j) enforce logical nonnegativity restrictions, and require that the variables x_{ijk} are fixed at the corresponding pre-specified values $\tilde{x}_{ijk}$ for the existing arcs $(i, j) \in A - P$ in the network. Note that by enforcing that pre-existing pipes be neither replaced, or retained without any parallel connections, Problem NOP reduces to the network design problem of Sherali *et al*. (1998).

Our principal set of decision variables are the lengths x_{ijk} of the different segments comprising each link $(i, j) \in P$, and the additional head H_{si} to be developed at each source node $i \in S$. The resulting heads H_i at each node $i \in N$ (above the elevation E_i of the node) and the flows q_{ij} in the links $(i, j) \in A$ are also problem variables that happen to be governed by the foregoing design variables.

5.4 Lower and Upper-bounding Problems

The frictional head loss expression in the Constraints (5.3b) cause NOP(W) to become nonlinear and nonconvex. Following Eiger *et al*. (1994) and Sherali *et al*. (1998), we take advantage of the monotone nature of these constraints to transform the problem NOP(W) into a set of newly

defined variables, and accordingly, develop suitable relaxations for the flow conservation constraints that turn out to be nonlinear in the projected space of these new decision variables.

Since the relations derived subsequently hold true for each link, the subscripts defining the links will be dropped for convenience. Equation (5.2) which appears in Constraints (5.3b) and can be written as follows using Equation (5.1), for any link having a flow q and a length segment vector $\boldsymbol{x} = (x_k, k = 1, ..., K)$.

$$F(q, \boldsymbol{x}) = \sum_{k=1}^{K} \text{sign}(q) \mid q \mid^{1.852} \ (1.52) \ 10^4 (C_{HW(k)})^{-1.852} d_k^{-4.87} x_k . \tag{5.4}$$

Denoting

$$v(q) \equiv \text{sign}(q) \mid q \mid^{1.852}, \tag{5.5}$$

Equation (5.4) can be rewritten as follows,

$$F(q, \boldsymbol{x}) = \sum_{k=1}^{K} v(q) a_k x_k , \tag{5.6}$$

where $a_k \equiv (1.52) \ 10^4 (C_{HW(k)})^{-1.852} d_k^{-4.87}$. (5.7)

By the monotonicity of $v(q)$, we can represent its value for any q as some convex combination of its minimum and maximum values $v(q\text{min})$ and $v(q\text{max})$, henceforth denoted as $v\text{min}$ and $v\text{max}$, respectively.

$$v(q) = l \ v\text{min} + (1-l) \ v\text{max}, \quad \text{for some } 0 \le l \le 1. \tag{5.8}$$

Using the representation (5.8) in Equation (5.6) and rearranging terms, we get,

$$F(q, \boldsymbol{x}) = \sum_{k=1}^{K} (v\text{min}) a_k (l x_k) + \sum_{k=1}^{K} (v\text{max}) a_k (1 - l) x_k . \tag{5.9}$$

We now define our new decision variables as

$$x_k^1 = l x_k \text{ and } x_k^2 = (1 - l) x_k , \tag{5.10}$$

so that

$$x_k = x_k^1 + x_k^2 . \tag{5.11}$$

Note that for the existing pipes, x_k is fixed at a value $\tilde{x}_k$ (some possibly zero) for all k. Equation (5.9) can now be rewritten in terms of the new decision variables as

$$F(q, \boldsymbol{x}) = \sum_{k=1}^{K} (v\text{min}) a_k x_k^1 + \sum_{k=1}^{K} (v\text{max}) a_k x_k^2 . \tag{5.12}$$

In the space of the new decision variables, x^1, x^2 and l, we have linearized the energy conservation constraints by substituting (5.12) on the left-hand side of (5.3b), but at the expense

of introducing nonlinearity elsewhere in the problem. Specifically, this nonlinearity arises in two sets of relationships. First, it occurs in the nonlinear representation (5.10) that accompanies the (linear) relationship (5.11). Second, the flow q (for each generic link) is now given via (5.5) and (5.8) by the following function $q(l)$,

$$q(l) = \text{sign}[l \ \ v\min + (1 - l)\ v\max] \mid l \ \ v\min + (1 - l)\ v\max \mid^{1/1.852}. \tag{5.13}$$

Using the foregoing transformations, we obtain an alternative equivalent representation for Problem NOP(W) as follows.

NOP(W):

$$\text{Minimize} \quad \sum_{(i,j)\in P} \sum_{k=1}^{K} c_k (x_{ijk}^1 + x_{ijk}^2) + \sum_{i\in S} c_{si} H_{si} \tag{5.14a}$$

subject to:

$$\sum_{k=1}^{K} (v\min_{ij}) \alpha_{ijk} x_{ijk}^1 + \sum_{k=1}^{K} (v\max_{ij}) \alpha_{ijk} x_{ijk}^2 = (H_i + E_i) - (H_j + E_j) \quad \forall\ (i, j) \in A \tag{5.14b}$$

$$\sum_{k=1}^{K} x_{ijk}^1 + \sum_{k=1}^{K} x_{ijk}^2 = L_{ij} \quad \forall\ (i, j) \in P \tag{5.14c}$$

$$\sum_{j\in FS(i)} q_{ij} - \sum_{j\in RS(i)} q_{ji} = b_i \quad \forall\ i \in D \tag{5.14d}$$

$$\sum_{j\in FS(i)} q_{ij} - \sum_{j\in RS(i)} q_{ji} \le b_i \quad \forall\ i \in S \tag{5.14e}$$

$$q_{ij} = \text{sign}[l_{ij}\ v\min_{ij} + (1 - l_{ij})\ v\max_{ij}] \mid l_{ij}\ v\min_{ij} + (1 - l_{ij})\ v\max_{ij} \mid^{1/1.852} \quad \forall\ (i, j) \in A \tag{5.14f}$$

$$q\min_{ij} \le q_{ij} \le q\max_{ij} \quad \forall\ (i, j) \in A \tag{5.14g}$$

$$H_i + E_i \le F_i + H_{si} \quad \forall\ i \in S \tag{5.14h}$$

$$H_{iL} \le H_i + E_i \le H_{iU} \quad \forall\ i \in D \tag{5.14i}$$

$$H_{si} \ge 0 \quad \forall\ i \in S \tag{5.14j}$$

$$x_{ijk}^1, x_{ijk}^2 \ge 0 \quad \forall\ (i, j) \in P,\ k = 1, \ldots, K \tag{5.14k}$$

$$0 \le l_{ij} \le 1 \qquad \forall\ (i, j) \in A \tag{5.14λ}$$

$$x^1_{ijk} = l_{ij}\tilde{x}_{ijk}, \text{ and } x^2_{ijk} = (1 - l_{ij})\tilde{x}_{ijk} \qquad \forall\ (i, j) \in A - P,\ k = 1, ..., K \tag{5.14m}$$

$$x^1_{ijk} = l_{ij}x_{ijk}, \text{ and } x^2_{ijk} = (1 - l_{ij})x_{ijk}, \text{ where } x_{ijk} \ge 0,\ \forall\ (i, j) \in P,\ k = 1, ..., K. \tag{5.14n}$$

Now we will construct relaxations for the nonlinear relationships (5.14f) and (5.14n) in order to derive lower-bounding linear programs.

5.4.1 Lower-bounding Relaxation LB(*W*): To derive this relaxation for Problem NOP(*W*) as given by (5.14), we first omit (5.14n), but replace it with its following aggregate relationship based on (5.3c):

$$\sum_{k=1}^{K} x^1_{ijk} = l_{ij}L_{ij} \quad \forall (i, j) \in P. \tag{5.15}$$

Note that the symmetric relationship $\sum_{k=1}^{K} x^2_{ijk} = (1 - l_{ij})L_{ij}\ \ \forall (i, j) \in P$ is implied by (5.15) and (5.14c). Next, we relax the nonlinear relationship (5.14f) by constructing a polyhedral outer-approximation to this function which relates q_{ij} to $l_{ij}\ \ \forall (i, j) \in A$ as in Sherali *et al.* (1998). In the most general case, this function is concave-convex as depicted in Figure 5.1 for the generic representation $q(l)$ stated as a function of l as in (5.13). In this figure, $\bar{l}$ is such that, if it exists, the tangential support to $q(l)$ at $l = \bar{l}$ passes through the coordinate (1, *q*min) in the (l, *q*) space. Similarly, $\tilde{l}$ is such that, if it exists, the tangential support to $q(l)$ at $l = \tilde{l}$ passes through the coordinate (0, *q*max) in the (l, *q*) space. These two values are respectively computed via Equations (5.16) and (5.17) given below using a bisection search, and each has a solution if and only if the corresponding entity $\bar{l}$ or $\tilde{l}$ exists.

$$q\text{min} - q(\bar{l}) - (1 - \bar{l})q'(\bar{l}) = 0 \qquad \text{for } 0 < \bar{l} < 1, \tag{5.16}$$

$$q\text{max} - q(\tilde{l}) + \tilde{l}q'(\tilde{l}) = 0 \qquad \text{for } 0 < \tilde{l} < 1, \tag{5.17}$$

where, the derivative (or slope) of $q(l)$, denoted by $q'(l)$, is given by

$$q'(l) = \frac{(v\text{min} - v\text{max})}{1.852} \,|\, l\, v\text{min} + (1 - l)v\text{max} \,|^{\frac{-0.852}{1.852}}, \text{ for } l \in [0,1], l\, v\text{min} + (1 - l)v\text{max} \ne 0.$$

To enhance the relaxation used in Sherali *et al.* (1998), we construct the following, more refined support structure. For each of the following cases, the corresponding nonlinear constraint (5.14f) is replaced by a set of affine inequalities that this relationship must satisfy in the (l, *q*)-space. For the general concave-convex case of Figure 5.1, if $\bar{l}$ and $\tilde{l}$ exist, we construct six supporting facets for the polyhedral approximation via tangents at {0, $\bar{l}/2$, $\bar{l}$} and at {$\tilde{l}$, $(1 + \tilde{l})/2$, 1} as shown in Figure 5.1. In case $\bar{l}$ does not exist, but $\tilde{l}$ does, or vice versa, we replace the

corresponding undefined supports with the affine concave or convex envelope, respectively, that pass through (0, $q$max) and (1, $q$min). This case yields only four supporting facets for the polyhedral approximation.

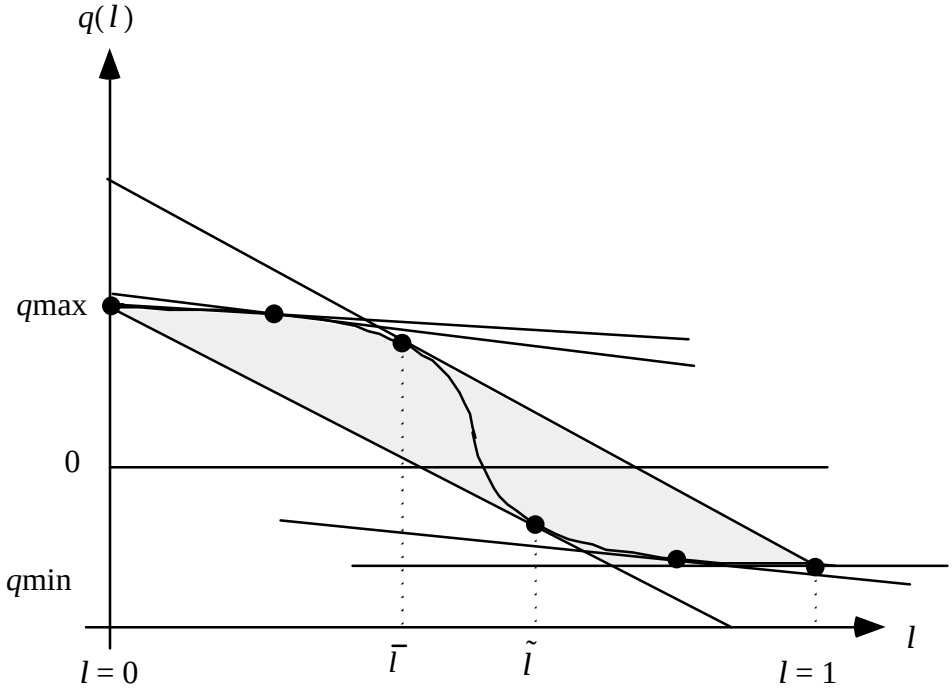

Figure 5.1. Polyhedral outer-approximation for relating $q(l)$ to l.

In case $q(l)$ is a concave function of l (i.e., $q$max > 0 and $q$min ≥ 0), we construct the affine convex envelope along with tangential supports at the points l = 0, 0.25, 0.5, 0.8, and 0.9. Similarly, if $q$max ≤ 0 and $q$min < 0, so that $q(l)$ is a convex function of l, we construct its affine concave envelope along with tangential supports at the points l = 0.1, 0.2, 0.5, 0.75, and 1. Each of these cases produce six facets for the polyhedral approximation that replaces (5.14f) in the relaxation LB(W). We also experimented with using four facets for the polyhedral approximation, instead of the six used here, similar to the consideration of Sherali *et al*. (1998). We provide some comparative computational experience for this in Chapter 6.

5.4.2 Lower-bounding Relaxation RLT(W**):** As an alternative to LB(W), we derive a further enhanced lower-bounding procedure that is motivated by the Reformulation-Linearization Technique (RLT) of Sherali and Tuncbilek (1992) for solving polynomial programming problems. The purpose of this report is to study the tradeoff between a quicker versus a more involved, but stronger, lower-bounding procedure with respect to the overall effort for solving the problem. To construct such a lower-bounding problem RLT(W), we augment Problem LB(W) by incorporating certain additional constraints that are generated using the RLT concept as follows.

Reformulation Step. The following quadratic valid constraints are generated based on the products of the stated pairs of inequalities (written in the form $\{\cdot\} \geq 0$), or based on the products of the equations with variables.

(a) Using the pipe length constraints in (5.14c), generate the equality product constraints

$$\left(\sum_{k=1}^{K} x_{ijk}^{1} + \sum_{k=1}^{K} x_{ijk}^{2}\right) q_{ij} = L_{ij} q_{ij} \quad \forall (i, j) \in P.$$

(b) Multiply each constraint in LB(W) that represents a linear inequality in q_{ij} and l_{ij} for the corresponding polyhedral approximation to (5.14f), with each corresponding variable x_{ijk}, $\forall (i, j) \in P,\ k = 1, ..., K$.

Linearization Step. Linearize the resulting product constraints generated above by substituting $y_{ijk} = q_{ij} x_{ijk}$, $\forall (i, j) \in P,\ k = 1, ..., K$, and by using equations (5.10) and (5.11). This produces a linear programming lower-bounding problem RLT(W) that incorporates certain additional valid inequalities that must be satisfied by any feasible solution to the nonlinear problem NOP(W) given by (5.14).

Note that if n_s is the number of supporting hyperplanes used to develop the polyhedral outer approximation for each arc flow, the RLT formulation has ($n_s|A|K+|A|$) additional constraints than the lower-bounding problem LB(W). In Chapter 6, we present some computational comparisons for the branch-and-bound procedure using these two lower-bounding schemes.

5.4.3 Upper Bounds: For computing upper bounds on the least-cost pipe sizing problem NOP(W), we fix the conserving flow solution as obtained via the lower-bounding problem LB(W) or RLT(W) within the Problem NOP(W), and solve the resulting linear programming problem. If this problem is feasible, it yields an upper-bounding completion to this fixed flow. As an alternative, a more refined, but computationally non-intensive local search heuristic could be employed to derive improved upper bounds.

5.5 A Branch-and-Bound Algorithm

We embed the lower and upper-bounding schemes described in the foregoing section in a branch-and-bound procedure to solve NOP(W) globally to any specified percentage tolerance (100e%) of optimality. Each branch-and-bound node principally differs in the specification of the hyperrectangle W. The hyperrectangle associated with node t of the branch-and-bound tree at the main iteration or stage S of the procedure is denoted by $W^{S,t} = \{\boldsymbol{q} : \boldsymbol{q}\mathbf{min}^{S,t} < \boldsymbol{q} < \boldsymbol{q}\mathbf{max}^{S,t}\}$. In our implementation of the branch-and-bound procedure, we successively partition the hyperrectangle defined by the initial bounds $W^{1,1} \equiv W$ on the flow variables into smaller and smaller hyperrectangles. At any stage S of the branch-and-bound algorithm, we have a set of active or nonfathomed nodes denoted as T_S. We select an active node t^* in T_S that has the least lower bound (this is termed as the *global lower bound* GLB_S at stage S), breaking ties arbitrarily, and partition the hyperrectangle associated with this node according to a suitable branching variable selection strategy that is described below. The selection of a branching variable according to this strategy ensures convergence of the overall procedure to a global optimum for NOP(W) using the general theory discussed in Sherali *et al*. (1998). This process continues by solving the bounding problems for the resulting two node subproblems, and then fathoming the nodes for which the lower bound is greater than or equal to UB(1 – e), where UB is the value of

the current incumbent solution, and $0 < e < 1$ is a suitable tolerance. ($e = 10^{-6}$ was used in the computations, thereby obtaining solutions to within 10^{-4} % of optimality.) Whenever the set of active nodes is empty, the process terminates. A formal statement of this algorithm is given in Sherali *et al*. (19988). Here, the focus on the principal difference was based on (a) a reduction in the potential set of branching variables, and (b) a strategy for selecting a branching variable from this set. These two features are described in succession below.

5.5.1 Maximal Spanning Tree-Based Approach for Reducing the Candidate Set of Branching Variables (MSTR): We can reduce the number of possible candidates for selecting a branching variable by restricting these to be a set of independent arcs in A. To see this, suppose that at the beginning of the branch-and-bound procedure, we construct a maximal spanning tree for the distribution network via Kruskal's (1956) procedure, using arc weights $(q\max_{ij} - q\min_{ij})\ \forall (i,\ j) \in A$. (The supply nodes are connected to a dummy sink via slack arcs having a large weight for this purpose, in order to balance supply and demand, and thereby obtain an equality flow conservation system. Hence, all these slack arcs are a part of the maximal spanning. Let B denote the set of arcs in this spanning tree. The remaining arcs $\{A–B\}$ in the network are designated as non-tree arcs and form the set from which the branching variables are selected. Note that such a spanning tree yields a valid basis for the underlying network flow problem, and given the flows on the independent nonbasic arcs, the corresponding flows on the dependent basic arcs are uniquely determined. Hence, only the nonbasic arcs corresponding to this (fixed) basis are selected for partitioning flow intervals. Upon fixing the flow bounds for the set of nonbasic arcs, the flow bounds for the basic arcs are updated using the representation of the dependent basic variables in terms of the independent nonbasic variables (see Bazaraa *et al*., 1990). Since most water distribution networks are almost "tree-like," this reduction in the potential set of branching variables is substantial.

Given the flow bounds on the nonbasic arcs, the basic flow bounds are updated following the reverse thread (post-order) recursive tree traversal procedure that is typically employed for computing flows in network simplex implementations (see Bazaraa *et al*. (1990), for example). At any step in this process, while examining a node j with the basic arc $(i,\ j) \in B$ leading into node j (node i being the predecessor of node j), we perform the following update operation (a similar operation is performed if the corresponding arc is $(j,\ i)$, leading out of node j):

$$q\max_{ij} = \min\{q\max_{ij}, -b_j - \sum_{\substack{k \in RS(j) \\ k \neq i}} q\min_{kj} + \sum_{k \in FS(j)} q\max_{jk}\} \tag{5.18}$$

$$q\min_{ij} = \max\{q\min_{ij}, -b_j - \sum_{\substack{k \in RS(j) \\ k \neq i}} q\max_{kj} + \sum_{k \in FS(j)} q\min_{jk}\} \tag{5.19}$$

Hence, as the flow interval lengths for nonbasic arcs shrink to zero, so do the corresponding interval lengths for the basic arcs since the right-hand sides of (5.18) and (5.19) coincide in this case at each step.

5.5.2 Branching Variable Selection Strategy: The recommended rule combines a pair of strategies based primarily on the fact that the polyhedral approximations in the $(q,\ l)$ space are

less exact when the flow is distant from both of its bounds. Hence, when partitioning any node of the branch-and-bound tree, the base-strategy is to select the branching variable (r, s) according to

$$(r, s) \in \text{argmax}\{\min\{q\max_{ij} - \hat{q}_{ij}, \hat{q}_{ij} - q\min_{ij}\} : (i, j) \in A - B\}. \tag{5.20}$$

where B is the set of tree arcs in the MSTR procedure described above, and $\hat{q}$ is the flow solution produced by solving the lower-bounding relaxation. The bounding intervals for q_{rs} in the two children node sub-problems are then taken as $[q\min_{rs}, \hat{q}_{rs}]$ and $[\hat{q}_{rs}, q\max_{rs}]$, respectively.

However, even if the non-basic flow values are driven close to either their lower or upper-bounding interval end-points as the algorithm progresses, the implied interval bounds on the basic arcs as determined by (5.18) and (5.19) are not sufficiently tight, and the same situation may not be the case for these basic arcs, thereby making the process stall using this strategy. Hence, when a substantial improvement in the global lower bound is not obtained for any pair of successive stages, i.e., if at any stage $S + 1$, it turns out that $0.9\,GLB_{S+1} \leq GLB_S$, then at this stage, we adopt an alternative branching variable selection strategy. As explained below, this also induces convergence to a global optimum.

The proposed switchover strategy is based on simply partitioning the longest flow interval among the nonbasic arcs, i.e., the branching variable index (r, s) is selected as

$$(r, s) = \underset{(i,j)}{\text{argmax}}\{q\max_{ij} - q\min_{ij} : (i, j) \in A - B\}. \tag{5.21}$$

In this case, having selected (r, s) according to (5.21), we bi-partition the interval $[q\min_{rs}, q\max_{rs}]$ by cutting it at the value 0 if $q\min_{rs} < 0 < q\max_{rs}$, or by bisecting this interval, otherwise. We found this combination strategy to be more effective than using either scheme on its own, or using the strategy prescribed in Sherali *et al*. (1998) (see Subramanian (1999) for detailed comparative results).

To see how this induces convergence, let us define d_{ij}, for each $(i, j) \in A$, as the discrepancy in the actual (Equation (5.3b)) versus the approximate (Equation (5.14b)) head-loss computation relative to the solution produced by the relaxed solution, and denote $D = \max\{d_{ij} : (i, j) \in A\}$. Note that if $D = 0$ (empirically measured by $D \leq 10^{-6}$ in our implementation), then we will have achieved an optimal solution to the current node sub-problem, because the solution $(\boldsymbol{q}, \boldsymbol{H_s}, \boldsymbol{H}, \boldsymbol{x})$ where $\boldsymbol{x} \equiv \boldsymbol{x}^1 + \boldsymbol{x}^2$, that is produced by the relaxation LB(W) (or RLT(W)) will be feasible to NOP(W) as defined in (5.3), yielding the same objective value. Hence, this node can be fathomed after updating the incumbent solution. Moreover, when $q\min_{ij} = q\max_{ij}$ $\forall (i, j) \in A - B$, we will also have $q\min_{ij} = q\max_{ij}$ $\forall (i, j) \in B$ by (7.1) and (7.2), and then as shown in Sherali *et al*. (1998), this will imply that the condition $D = 0$ holds true, thereby indicating optimality for the current subproblem. Hence, if any path of the branch-and-bound tree associated with selecting the node having the least lower bound at each stage is examined, the condition $0.9\ GLB_{S+1} > GLB_S$ can hold only finitiely often along this path. This induces a process whereby $\{q\max_{ij} - q\min_{ij}\} \to 0$ $\forall (i, j) \in A - B$, leading to $D \to 0$ as above along such a path. Therefore, a global optimum is recovered in the limit.

5.6 Summary

In this chapter, we have proposed enhanced lower-bounding procedures along with significantly more effective branching and partitioning strategies for determining global optimal solutions to water distribution network design problems. The application of efficient schemes such as those described in Sherali and Smith (1997) to obtain tight flow bounds or upper bounds for each node subproblem is more critical in the case of the network design problems having several independent variables. Such problems can also benefit via the construction of tighter lower-bounding problems through the use of an additional, suitable number of supporting hyperplanes in the approximation of the flow relationships, once the direction of flow in any link is determined, as well as through the proposed RLT constructs. Such investigations and further computational tests are proposed for future research.

CHAPTER 6

APPLICATION TO EXAMPLE NETWORKS

6.1 Introduction

In this chapter, we apply the proposed branch-and-bound algorithm to three standard test problems from the literature and a newly generated Blacksburg network. The algorithms were implemented on a SUN SPARC 10 UNIX workstation, using the CPLEX 6.0 callable library to solve the linear programming problems. The computer code was written in C++. The algorithm was implemented using a termination criterion of $e = 10^{-6}$. In addition, we also experimented with using a lesser number of supporting hyperplanes (four) in lieu of six as discussed in Chapter 5 for constructing the lower-bounding linear programs, with and without RLT enhancements. The computational results and the best design configuration for each of these four networks are presented sequentially below.

6.2 Test Problems

Test Problem 1: Two-Loop Network: This is a single source test problem originally presented by Alperovits and Shamir (1977). (See Sherali *et al*. (1998) for a description of the network configuration and for the set of initial bounds on the flows. The latter were logically determined as previously stated in Chapter 5). The set of commercially available pipes were taken as having diameters d (inches) given by {1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24}, with the corresponding costs per unit length ($/meter) being {2, 5, 8, 11, 16, 23, 32, 50, 60, 90, 130, 170, 300, 550}. Note that the original Alperovits and Shamir (1977) test problem excludes certain pipe diameters, whereas several authors have later solved this problem by including all the possible aforementioned diameters. To enable a comparison, as well as from practical viewpoint, we permit the selection of all commercially available pipe diameters. Table 6.1a summarizes the results obtained using different numbers of supports (n_s) per arc, and Table 6.1b provides the key information regarding the best design obtained.

Table 6.1a. Computational Results for the Two-Loop Network.

n_s	Global Lower Bound	Global Upper Bound	# LP Solved	CPU Time (sec)
4	403385.1	403385.401	141	11.64
6	403385.093	403385.408	147	18.05

Table 6.1b. Optimum Design for the Two-Loop Network ($e = 10^{-6}$).

Arc #	Segments having Length (m)	Flow (m^3/hr)	Diameter (inches)	Head Loss (m)
1	1000	1120.0	18	6.749242
2	795.408540	368.331588	10	11.982436
2	204.591460	368.331588	12	1.268322
3	1000	651.668412	16	4.393325
4	999.998776	0.975686	1	18.857409
4	0.001224	0.975686	2	0.000001
5	310.354145	530.692726	14	1.786092
5	689.645855	530.692726	16	2.071341
6	11.139149	200.692726	8	0.161574
6	988.860851	200.692726	10	4.838426
7	98.493861	268.331588	8	2.446478
7	901.506139	268.331588	10	7.553516
8	1000	–0.692726	1	–10.000004

Among the more recent results of this problem, the heuristic of Loganathan *et al*. (1995) found a solution having a total cost of $403,657. Sherali *et al*. (1998) obtained a somewhat improved solution with an objective value of $403,390. Earlier, Sherali and Smith (1995) had obtained a global lower bound of $403,385 on this problem, along with a feasible solution of $403,386, which is $1 within global optimality. Their algorithm, when implemented on the same computer and using CPLEX 2.0 to solve the LP relaxations, enumerated only 49 nodes, but consumed 342 CPU seconds due to the size of their lower-bounding problem. The best solution presented in Table 6.1b is obtained using our branch-and-bound procedure with four supporting hyperplanes per arc. In the literature, this solution is the most accurate as reported, and has an objective value that lies within $0.2 of global optimality, and was derived while consuming only 12 CPU seconds. The use of the procedure, MSTR, resulted in a reduction in computational effort by a factor of 5, hence underscoring the usefulness of this scheme.

Test Problem 2: Hanoi Network: The Hanoi network is a single source network consisting of 3 basic loops, 32 nodes and 34 links. The network configuration, arc definitions, and flow bounds are given in Sherali *et al*. (1998), and the other data appears in Fujiwara and Khang (1990).

Table 6.2a presents the results obtained using our algorithm, and Table 6.2b provides the key design parameters for the best solution reported.

Table 6.2a. Computational Results for the Hanoi Network.

n_s	Global Lower Bound	Global Upper Bound	# LP Solved	CPU Time
4	6055536.43	6055542.48	1533	245.44 sec
6	6055536.31	6055542.37	1075	552.82 sec

Eiger *et al*. (1994) reported a solution having an objective function value of $6,026,660 for this problem using an optimality tolerance of 0.5%. However, their solution contains some violations in the flow conservation constraints, as shown by Sherali *et al*. (1998), who obtained a solution having an objective value of $6,058,976, which is the best solution previously reported in the literature. The best solution found by our algorithm has an objective value of $6,055,542 which is significantly better than the values reported in the literature for this test problem. This solution was obtained within 4 minutes of CPU time, using 4 supporting hyperplanes per arc in LB(W), and is within 10^{-4} % of optimality (or $6 of global optimality) as verified by our global lower bound.

The original data for the Hanoi test network presented in Fujiwara and Khang (1990) used a C_{HW} value of 162.5. For the sake of comparison, the above run was repeated using this value for the Hazen-William coefficient. A global lower bound of 4,954,941.69 and a corresponding feasible solution having an objective cost of 4,954,945.29 was obtained after solving 541 linear programs and expending 2 minutes of CPU time. This solution is within $4 of global optimality and significantly improves upon the best objective cost of 5,562,000 reported in Fujiwara and Khang (1990).

Table 6.2b. Optimum Design for the Hanoi Network ($e = 10^{-6}$).

Arc #	Dia	Length (m)	Flow (m^3/hr	Head Loss (m)	Arc #	Dia	Length (m)	Flow (m^3/hr	Head Loss (m)
1	40	100	19940	2.859823	19	24	400	2403.036233	2.734511
2	40	1350	19050	35.477043	20	40	2200	7831.762902	11.145477
3	40	900	7965.200865	4.704430	21	16	491.360368	1415	9.074382
4	40	1150	7835.200865	5.830782	21	20	1008.639632	1415	6.283494
5	40	1450	7110.200865	6.141881	22	12	500	485	5.159782
6	40	450	6105.200865	1.437398	23	40	2650	5141.762902	6.158480
7	40	850	4755.200865	1.709164	24	30	1230	3501.070442	5.694622
8	40	850	4205.200865	1.361194	25	30	1300	2681.070442	3.671724
9	30	74.233811	3680.200865	0.376960	26	20	850	1186.762902	3.823008
9	40	725.766189	3680.200865	0.907902	27	12	299.999150	286.762902	1.169823
10	30	950	2000	1.559305	27	16	0.000850	286.762902	0.000001
11	24	1200	1500	3.427313	28	12	750	83.237098	0.295907
12	24	3500	940	4.206805	29	16	1500	595.692460	5.580180
13	16	253.562337	1155.200865	3.216195	30	12	1999.999879	305.692460	8.778997
13	20	546.437663	1155.200865	2.338011	30	16	0.000121	305.692460	0
14	16	500	540.200865	1.551952	31	12	1600	–54.307540	–0.286252
15	12	550	260.200865	1.791357	32	16	150	–414.307540	–0.284832
16	12	2730	–133.036233	–2.566981	33	16	748.166501	–519.307540	–2.158642
17	16	1750	–998.036233	–16.930625	33	20	111.833499	–519.307540	–0.108844
18	20	419.803213	–2343.036233	–6.654887	34	24	950	1324.307540	2.154262
18	24	380.196787	–2343.036233	–2.480223					

Test Problem 3: New York Network: The New York test network configuration and data are given in Loganathan *et al*. (1985) and Fujiwara and Khang (1990). The original network has 20 nodes and 21 arcs, while the expanded network (with parallel arcs) has 26 nodes and 33 arcs (see Subramanian (1999) for details). Since there is only a single source node, the initial flow bounds for the arcs were calculated using the procedure described in Sherali and Smith (1997). The

computational results and the best design obtained are presented in Tables 6.3a and 6.3b, respectively. The coefficients used in the head-loss equation are the same as that used in Fujiwara and Khang (1990) and Loganathan *et al*. (1985). The flow rate exponent value was set equal to 1.85, while the head loss coefficient was set equal to 851500 to conform with the flow rates being measured in cubic-feet per second, the pipe diameters in inches, and the head losses in feet.

Table 6.3a. Computational Results for the New York Network ($e = 10^{-6}$).

n_s	Global Lower Bound	Global Upper Bound	# LP Solved	CPU Time
4	37878580.89	37878581.28	12953	93 min
6	37878580.89	37878581.28	10957	58 min

The New York test network problem was first analyzed using parallel links by Schaake and Lai (1969) and they obtained a solution having an objective value of $77.61($10^6$). Fujiwara and Khang (1990) applied their two-phase approach to this problem, but the solution presented by them was not feasible. Quindry *et al*. (1981) obtained a solution having a total cost of $63.581($10^6$), while Gessler (1982), Bhave (1985), and Morgan and Goulter (1985) obtained solutions having costs of $41.2($10^6$), $40.18($10^6$), and $39.018($10^6$), respectively. Loganathan *et al*. (1995) used a simulated annealing-based heuristic procedure to further improve the objective value to $38.04($10^6$). All these approaches are heuristic in nature and simply seek to determine (at best) local optimal solutions, providing no indication of a competitive global lower bound on the optimum value. Using our procedure, we were able to obtain such a global lower bound of value of $37,878,580.89 and a feasible solution having a cost of $37,878.581.28. This is the best solution reported thus far in the literature and lies within 10^{-6} % (or within $0.4) of optimality. Note that this result is obtained by considering 4-inch pipe increments for the set of available pipe diameters. Since 1995, all the results presented above use 12-inch pipe diameter increments for the sake of computational ease. Hence, for the sake of comparison, our procedure (using six supports per arc) was run using 12-inch pipe increments. The results produced a global lower bound value of $38,067,895 along with a corresponding feasible solution of $38,067,935, after solving 2467 linear programs. It was observed that while the optimal pipe diameters coincided with that obtained by Loganathan *et al*. (1995), the corresponding pipe segments lengths were different. This is due to the fact that the head loss values obtained by Loganathan *et al*. (1995) have a feasibility tolerance of 0.1, while the results presented in this paper are obtained using a more precise feasibility tolerance value of 10^{-6}. Consequently, the objective cost corresponding to the optimal solution is higher than that obtained by Loganathan *et al*. (1995), and represents a relatively more accurate estimate of the actual solution.

It can be seen from Table 6.3a that the computational times are significantly higher for this test case, as compared with the computational efforts for the previous two test problems. One important reason for this difference is that no logical test based schemes were used to generate tight initial flow interval bounds. In fact, the initial feasible solution was itself near-optimal, but was polished to the final solution only toward the tail-end of the branching procedure. The results for this test problem also differ from the previous two in that the introduction of

additional hyperplanes results in an improvement in the computational effort, both in terms of the number of nodes enumerated and the CPU time expense.

Test Problem 4: Blacksburg Test Network: Figure 6.1 depicts a network representation of a newly expanded subdivision of the water distribution system in the town of Blacksburg, Virginia. The network data for this problem was acquired from the public works department of the town, along with other problem parameters such as pressure requirements, locations of fire hydrants, cost factors, pipe quality that is reflected via the associated C_{HW} value, and demand requirements. The unit pipe costs used in this problem represent real-life values and include installation costs as well. The link and node data for the network are presented in Tables 6.4a and 6.4b. The set of pipes whose diameters are fixed is listed in Table 6.4a. A Hazen-Williams coefficient value of 120 was used for all the links. Expression (2.1) was used to compute the head losses. The flow rates were converted using double precision from gallons per minute (gpm) into units of m^3/hr, the pipe diameters were specified in centimeters, and the head losses in meters. The computational results and the best design obtained are presented in Tables 6.4c and 6.4d, respectively. The optimal flow values can be computed using the optimal head loss and pipe diameter values given in Table 6.4d via Equation (2.1) (using appropriate coefficients for the measurement units specified above).

The computational results for the Blacksburg network also exhibit that the introduction of additional supporting hyperplanes (six instead of four per arc), results in a decrease in the number of nodes enumerated, as well as in the CPU time expended. The best solution obtained for this test problem has an objective value of 577066, along with a best global lower bound of 577066.

6.3 Discussion on Algorithmic Strategies and the Use of RLT(W).

The computational experience on the foregoing test problems clearly indicated that the maximal spanning tree reduction procedure (MSTR) is an indispensable strategy. In fact, when we suppressed the procedure MSTR for the two larger New York and Blacksburg test problems, we were unable to obtain good quality feasible solutions within the time limit of 10 CPU hours. Furthermore, the use of six versus four hyperplanes per arc in the polyhedral approximation of the flow relationships improved the relative performance for these two larger test problems, but it resulted in only a marginally greater CPU time expense for the other two test problems,.

However, in the cases where the introduction of additional supports was beneficial, a significant reduction in computational time was observed. Hence, we recommend the use of the six prescribed supporting hyperplanes per arc in the lower-bounding problem, but suggest experimenting with additional supports.

Table 6.3b. Optimum Design for the New York Network ($e = 10^{-6}$).

Link Index	Dia (inches)	Segment Length (ft)	Flow (gpm)	New Cost ($)	Head Loss (ft)
1	180	11600	879.302755	7987090.046769	5.729565
2	180	19800	786.902755	13633136.459141	7.963917
3	180	7300	694.502755	5026358.391501	2.330383
4	180	8300	606.302755	5714900.636912	2.060921
5	180	8600	518.102755	5921463.310536	1.596524
6	180	19100	429.902755	13151156.887353	2.510590
71	132	4800	201.677399	2249808.727816	0.704444
72	132	4800	201.677399	2249808.727816	0.704444
7	112	2408.899652	140.025356	920962.065953	0.400677
7	116	7191.100348	140.025356	2871547.079089	1.008212
8	132	12500	253.502755	5858876.895354	2.800706
9	180	9600	58.500000	6610005.555947	0.031515
10	204	11200	131.124553	9006418.334910	0.088965
11	204	14500	484.997245	11660095.165731	1.295005
12	204	12200	836.297245	9810562.829098	2.985471
13	204	24100	953.397245	19379882.309939	7.515528
14	204	21100	1045.797245	16967448.827374	7.808112
15	204	15500	1138.197245	12464239.659919	6.708414
161	72	13200	17.039827	2917822.789095	0.383495
162	72	13200	17.039827	2917822.789095	0.383495
16	100	26399.999870	40.460173	8769939.752862	0.766990
16	104	0.000130	40.460173	0.045337	0.000000
171	72	15600	69.403955	3448336.023476	6.090610
172	72	15600	69.403955	3448336.023476	6.090610
17	96	0.000063	164.796045	0.019895	0.000000
17	100	31199.999937	164.796045	10364474.283491	12.181220
181	60	12000	38.897986	2115833.464170	3.900627
182	60	12000	38.897986	2115833.464170	3.900627
18	76	9745.580952	78.202014	2303611.685506	3.646560
18	80	14254.419048	78.202014	3590655.926313	4.154695
191	60	7200	64.220140	1269500.078502	5.917147
192	60	7200	64.220140	1269500.078502	5.917147
19	72	0.195798	119.652552	43.280596	0.000209
19	76	14399.804202	119.652552	3403753.700483	11.834085
20	60	38400	13.872692	6770667.085344	1.853176
211	72	13200	80.619266	2917822.789095	6.799253
212	72	13200	80.619266	2917822.789095	6.799253
21	68	10595.684537	75.508042	2181889.521376	6.386791
21	72	15804.315463	75.508042	3493499.380606	7.211714

Total Cost	Existing Cost	New Cost
217,700,927	179,822,346	37,878,581

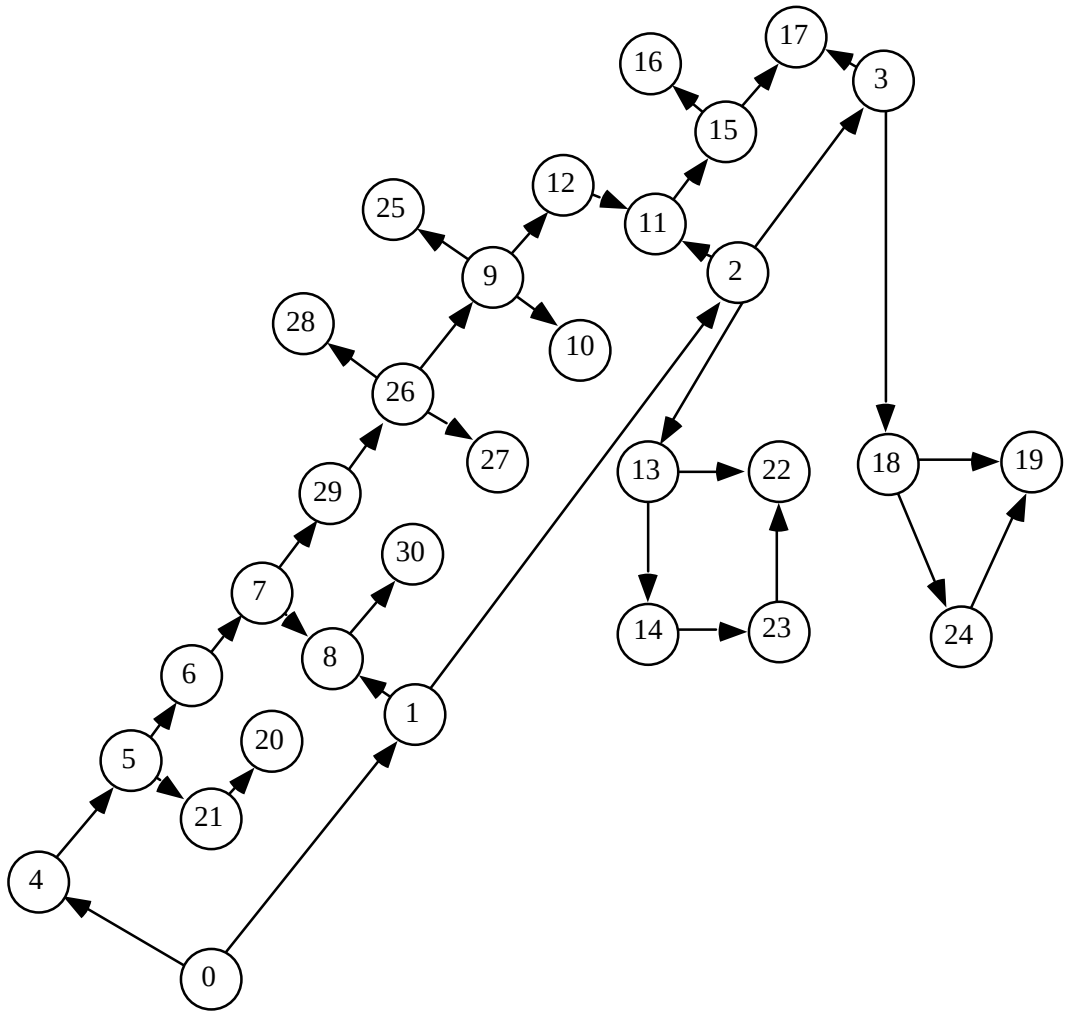

Figure 6.1. Blacksburg Test Network Configuration.

Table 6.4a. Arc Data for the Blacksburg Network.

Arc	Length (ft)	Fix Dia (inches)	Arc	Length (ft)	Fix Dia (inches)	Arc	Length (ft)	Fix Dia (inches)
(0, 1)	1363	—	(6, 7)	95	—	(15, 17)	1009	—
(0, 4)	194	—	(7, 8)	419	—	(18, 19)	408	—
(1, 2)	1832	—	(7, 29)	208	—	(18, 24)	1181	—
(1, 8)	151	—	(8, 30)	110	—	(21, 20)	113	—
(2, 3)	888	—	(9, 10)	451	6	(22, 13)	701	6
(2, 11)	155	—	(9, 12)	59	—	(23, 22)	351	6
(2, 13)	309	—	(9, 25)	416	6	(24, 19)	967	6
(3, 17)	699	—	(11, 15)	303	12	(26, 9)	271	—
(3, 18)	1151	—	(12, 11)	823	—	(26, 27)	317	6
(4, 5)	1098	—	(13, 14)	766	8	(26, 28)	424	6
(5, 6)	578	—	(14, 23)	382	—	(29, 26)	730	6
(5, 21)	611	6	(15, 16)	758	10			—

Table 6.4b. Node Data for the Blacksburg Network.

Node Index	Supply or Demand (gpm)	Elevation (ft)	Node Index	Supply or Demand (gpm)	Elevation (ft)
0	1548.63	2163	16	–52.11	2149
1	–52.11	2141	17	–10.38	2109
2	–50.58	2132	18	–103.65	2144
3	–25.77	2121	19	–52.11	2149.5
4	–13.84	2153.5	20	–10.96	2140
5	–53.65	2141.5	21	–51.35	2141.5
6	–51.73	2129	22	–11.73	2144
7	–200.58	2127	23	–51.54	2156.5
8	–11.35	2127	24	–102.50	2178
9	–10.77	2109.5	25	–51.54	2118
10	–52.11	2121	26	–52.50	2099.5
11	–100.77	2139	27	–50.96	2102
12	–27.11	2110	28	–25.58	2098.5
13	–100.77	2136.5	29	–41.92	2120
14	–25.77	2143	30	–51.35	2123
15	–51.54	2144.5			

Table 6.4c. Computational Results for the Blacksburg Network.

n_s	Global Lower Bound	Global Upper Bound	# LP Solved	CPU Time
4	577066	577067	2089	27 min
6	577066	577067	1975	24 min

Table 6.4d. Optimum Design for the Blacksburg Network ($e = 10^{-6}$).

Arc	Dia (inches)	Length (ft)	Arc	Dia (inches)	Length (ft
(13, 14)	8.0	766	(18, 24)	10.0	0.001887
(23, 22)	6.0	351	(18, 24)	12.0	1180.998113
(26, 28)	6.0	424	(7, 29)	10.0	208
(14, 23)	4.0	63.645446	(9, 25)	6.0	416
(14, 23)	6.0	318.354554	(9, 10)	6.0	451
(9, 12)	10.0	59	(26, 9)	6.0	92.889470
(2, 3)	16.0	888	(26, 9)	8.0	178.110530
(3, 18)	16.0	1151	(22, 13)	6.0	701
(12, 11)	12.0	823	(0, 4)	16.0	194
(11, 15)	12.0	303	(21, 20)	2.0	12.848676
(15, 16)	10.0	758	(21, 20)	3.0	100.151324
(0, 1)	24.0	1363	(1, 8)	3.0	43.747654
(2, 11)	12.0	69.689856	(1, 8)	4.0	107.252346
(2, 11)	16.0	85.310144	(5, 6)	12.0	577.999064
(2, 13)	10.0	309	(5, 6)	16.0	0.000936
(24, 19)	6.0	967	(26, 27)	6.0	317
(8, 30)	4.0	0.000108	(29, 26)	6.0	730
(8, 30)	6.0	109.999892	(4, 5)	16.0	1098
(1, 2)	20.0	13.727890	(5, 21)	6.0	611
(1, 2)	24.0	1818.272110	(3, 17)	1.0	698.988391
(15, 17)	3.0	714.434858	(3, 17)	2.0	0.011609
(15, 17)	4.0	294.565142	(6, 7)	12.0	95
(18, 19)	10.0	408	(7, 8)	4.0	419

Table 6.4d (continued). Optimum Design for the Blacksburg Network ($e = 10^{-6}$).

Node #	Head (ft)	Node #	Head (ft)
0	184.660000	16	93.308922
1	168.722421	17	46.160000
2	134.430362	18	90.189925
3	129.949917	19	80.263327
4	185.105409	20	46.160000
5	148.536939	21	72.661969
6	99.580936	22	66.147108
7	94.038161	23	46.160000
8	51.706685	24	46.160000
9	97.880803	25	53.029402
10	46.160000	26	75.785923
11	112.829253	27	46.160000
12	102.771162	28	66.662345
13	104.838176	29	97.379870
14	80.826940	30	46.160000
15	103.426382		

Table 6.5a. Comparative Results Between Using LB(W) Versus RLT(W) for $e = 10^{-3}$.

Network	**Nodes (RLT(W))/ Nodes (LB(W))**	**Time (RLT(W))/ Time (LB(W))**
Two-Loop	97/99 = 0.98	33/8 = 12.5
Hanoi	115/167 = 0.69	167/51 = 3.27
New York	6871/13673 = 0.50	255/71 = 3.59
Blacksburg	2734/3445 = 0.81	51/12 = 4.25

Table 6.5b. Comparative Results Between Using LB(W) Versus RLT(W) Without the Procedure MSTR for $e = 10^{-3}$.

Network	**Nodes (RLT(W))/ Nodes (LB(W))**	**Time (RLT(W))/ Time (LB(W))**
Two-Loop	521/887 = 0.59	109/61 = 1.79
Hanoi	2905/15681 = 0.19	49/65 = 0.75

As far as the use of the lower-bounding scheme RLT(Ω) is concerned, it was generally observed that this yielded much tighter lower bounds, and resulted in fewer branch-and-bound nodes being enumerated as compared with using LB(Ω). On the other hand, the compromise with respect to the computational effort was not favorable for these test cases, although as problem size increased, the relative benefit of using the tighter lower-bounding formulation RLT(Ω) became more pronounced. Table 6.5a illustrates this phenomenon. However, when we switched off the strategy MSTR, as seen from Table 6.5b for the Hanoi network, the RLT scheme enumerated significantly fewer nodes and also consumed less CPU time, in comparison to the relaxation LB(Ω). In general, it is apparent that the relative computational efficiency of the RLT-enhanced scheme improves as the size of the problem ($|A|$, $|N|$, K) and the number of possible choices for selecting the branching variable increase. This suggests that problems having several more independent variables than those analyzed in this paper, such as an RLT-enhanced procedure, might be beneficial. This is open to further investigation.

6.4 Summary

A new test problem dealing with the water distribution system in Blacksburg, Virginia is introduced to the literature in this chapter. Results obtained on this problem, as well as three other standard test problems from the literature, demonstrate the efficacy of the proposed methodology. Improved solutions are reported for each of the latter problems, significantly so for the two larger cases of the Hanoi and New York test networks for which solutions proven to lie within 10^{-4}% of optimality are derived for the first time in the literature. Further enhancements in algorithmic efficiency can be achieved by including a more effective preprocessor to deduce valid, tighter initial bounds on the flow variables. The algorithm can also benefit by computing sharper upper bounds by using some local optimization scheme, rather than simply evaluating the flow solution produced by the lower-bounding problem.

CHAPTER 7

SUMMARY

The main contribution of this research consists of the development of new methodologies for the optimal replacement of an individual water main and the global optimal design of an expanding water system. Because the proposed optimal threshold break rate is strictly analytical and provides an economically sustainable critical break rate, it should be of help in prioritizing failure-prone water mains for replacement. Also, the connection to the functions of reliability theory, such as the rate of occurrence of failure (ROCOF) and the hazard rate, are established. The design aids in the form graphs (Figure 3.2 and 3.3) permit engineers to better understand the system-wide failure pattern of pipes. The utility of the parametric form of the ROCOF function, the Weibull proportional intensity function capable of incorporating environmental variables is illustrated. The procedure also accommodates time-truncated data and a full failure history is not needed.

The global optimal design formulation presented in this report provides a holistic approach for modeling an expanding pipe network. As opposed to the techniques that search for improving local optima, the present formulation provides theoretical limits and is capable of reaching the global optimum. The implementation of the procedures is illustrated via examples. In combination, the critical threshold break rate allows for the progressive replacement of pipes and the optimal design scheme permits the suitable selection of pipe diameters.

REFERENCES

Agbenowosi, N., *A Mechanistic Analysis based Decision Support System for Scheduling Optimal Pipeline Replacement*, Ph.D dissertation, Virginia Tech, Blacksburg, VA, 2000.

Alperovits, E., and Shamir, U., "Design of Optimal Water Distribution Systems," *Water Resources Research*, Vol. 13, 1977, 885-900.

Andreou, S. A., D. H. Marks, and R. M. Clark, "A new methodology for modeling break failure patterns in deteriorating water distribution systems: Theory", *Advances in Water Resources*, vol.10, 2-10, 1987.

Andreou, S. A., D. H. Marks, and R. M. Clark. A new methodology for modeling break failure patterns in deteriorating water distribution systems: Applications, *Advances in Water Resources*, 10, 11-20, 1987.

Ascher, H. and H. Feingold, *Repairable Systems Reliability*, Marcel Dekker, Inc., New York, NY, 1984.

Bain, L.J., and Engelhardt, M., *Statistical Analysis of Reliability and Life-Testing Models*, Marcel Dekker, Inc., New York, NY, 1991.

Basalo, C., *Water and Gas Mains Corrosion, Degradation and Protection*, Ellis Horwood, Chichester, England, 1992.

Bazaraa, M.S., Jarvis, J. J., and Sherali, H.D., *Linear Program*min*g and Network Flows*, second edition, John Wiley and Sons, New York, NY, 1990.

Bhave, P.R., "Optimization of Gravity-Fed Water Distribution Systems: Theory," *Journal of Environmental Engineering*, ASCE, Vol. 109, No. 1, 1983, 189-205.

Clark, R. M., C. L. Stafford, and J. A. Goodrich, "Water distribution systems: A spatial and cost evaluation," *Journal of the Water Resources Planning and Management Division*, ASCE, 108, 243-256, 1982.

Dandy, G.C., Simpson, A.R., and Murphy, L.J., "An Improved Genetic Algorithm for Pipe Network Optimization," *Water Resources Research*, Vol. 32, No. 2, 1996, 449-458.

Deb, A.K., Y. Hasit, F. Grablutz, and R. Herz., *Quantifying Future Rehabilitation and Replacement Needs of Water Mains*. Denver, Colo.: AWWARF, 1998.

Eiger, G., Shamir, U., and Ben-Tal, A., "Optimal Design of Water Distribution Networks," *Water Resources Research*, Vol. 30, No. 9, 1994, 2637-2646.

Fujiwara, O., and Khang, D.B., “A Two-Phase Decomposition Method for Optimal Design of Looped Water Distribution Networks,” *Water Resources Research*, Vol. 26, No. 4, April 1990, 539-549.

Kettler, A. J. and Goulter, I. C. "An Analysis of Pipe Breakage in Urban Water Distribution Networks”, *Canadian Journal of Civil Engineering*, vol.12, pp.286-291, 1985.

Kleiner, Y., B. and Rajani, “Using limited data to assess future needs”, *Journal of the American Water Works Association*, Vol.91, No.7, pp.47-61, 1999.

Kleiner, Y., B. J. Adams, and J. S. Rogers, “Long-term planning methodology for water distribution system rehabilitation”, *Water Resources Research*, vol.34, pp.2039-2051, 1998.

Kleiner, Y., B. J. Adams, and J. S. Rogers, Selecting and scheduling rehabilitation alternatives, *Water Resources Research*, 34, 2053-2061, 1998.

Kruskal, J, B. “On the Shortest Spanning Tree of a Graph and the Traveling Salesman Problem,” *Proceedings of the American Mathematical Society*, Vol. 7, 1956, 48-50.

Lansey, K., and Mays, L., “A Methodology for Optimal Network Design,” *Computer Applications in Water Resources*, ed. H. C. Torno, 1985, 732-738.

Loganathan, G.V., Greene, J.J., and Ahn, T.J., “Design Heuristic for Globally Minimum Cost Water-Distribution Systems,” *Journal of Water Resources Planning and Management*, Vol. 121, No. 2, 1995, 182-192.

Male, J. W., T. M. Walski, and A. H. Slutski, Analyzing Water Main Replacement Policies, *Journal of Water Resources Planning and Management*, ASCE, 116, 363-374, 1990.

Mavin,K., "Predicting the Failure Performance of Individual Water Mains", Research Report No. 114, Urban Water Research Association of Australia, July 1996, ISBN 1876088176.

Mays, L., *Water Distribution Systems Handbook*, McGraw-Hill, 2000.

Morris, R.E., Jr., “The Distribution System”, in Manual of Water Utility Operations, Texas Water Utilities Association, Austin, Texas, 1975.

O’Day, D.K., Weiss,R., Chiavari, S., and Blair, D., *Guidance Manual: Water Main Evaluation for Rehabilitation / Replacement*, AWWA Research Foundation, Denver, CO, 1986.

Park, S., *An Optimal Pipeline Replacement Scheduling Model for Water Distribution Systems*, Ph.D dissertation, Virginia Tech, Blacksburg, VA, 2000.

Roberge, P.R., *Handbook of Corrosion Engineering*, McGraw-Hill, 2000.

Shamir, U., and C. D. D. Howard, An analytic approach to scheduling pipe replacement, *Journal of the American water Works Association* , vol.71, 248-258, 1979.

Sherali, H.D., and C.H. Tuncbilek, “A Global Optimization Algorithm for Polynomial Programming Problems Using a Reformulation-Linearization Technique,” *Journal of Global Optimization*, Vol. 2, 1992, 101-112.

Sherali, H.D., and Smith, E.P., “A Global Optimization Approach to a Water Distribution Network Design Problem,” *Journal of Global Optimization*, Vol. 11, 1997, 107-132.

Sherali, H.D., and Smith, E.P., “An Optimal Replacement-Design Model for a Reliable Water Distribution Network System,” *Integrated Computer Applications in Water Supply*, Vol. 1, Research Studies Press Ltd., Somerset, England, 1993, 61-75.

Sherali, H.D., Totlani, R., and Loganathan, G.V., “Enhanced Lower Bounds for the Global Optimization of Water Distribution Networks,” *Water Resources Research*, Vol. 34, No. 7, 1998, 1831-1841.

Subramanian, S. “Optimization Models and Analysis of Routing, Location, Distribution, and Design Problems on Networks.” Doctoral Dissertation, Virginia Polytechnic Institute and State University, 1999.

Walski, T. M., “Replacement rules for water mains”, *Journal of the American Water Works Association*, Vol. 79, pp.33-37, 1987.

Walski, T. M., and A. Pelliccia, "Economic Analysis of Water Main Breaks." *Journal of the American water Works Association*, 74(3), 140 – 147 , March 1982.

Walski, T.M., *Analysis of Water Distribution Systems*, Van Nostrand Reinhold Company, New York, NY, 1984.