R86-18

RELIABILITY OF WATER DISTRIBUTION SYSTEMS

BY
JANET M. WAGNER
URI SHAMIR
DAVID H. MARKS

RALPH M. PARSONS LABORATORY
HYDROLOGY AND WATER RESOURCE SYSTEMS

Report Number 312

Prepared under the support of the
National Science Foundation

JUNE 1986

MIT

DEPARTMENT
OF
CIVIL
ENGINEERING

SCHOOL OF ENGINEERING
MASSACHUSETTS INSTITUTE OF TECHNOLOG
Cambridge, Massachusetts 02139

R 86-18

RELIABILITY OF WATER DISTRIBUTION SYSTEMS

by

Janet M. Wagner

Uri Shamir

David H. Marks

RALPH M. PARSONS LABORATORY

HYDROLOGY AND WATER RESOURCE SYSTEMS

Report Number 312

Prepared under the support of the

National Science Foundation

JUNE 1986

RELIABILITY OF WATER DISTRIBUTION SYSTEMS

by

Janet M. Wagner
Uri Shamir
David H. Marks

ABSTRACT

Reliability of water distribution systems is becoming of increasing concern to water system designers and operators because of the increasing age of many systems and the decreasing availability of public money for water system construction and repair. This report is concerned with identifying, developing, and applying methods for the calculation of probabilistic reliability measures for water distribution systems. Methods are developed for assessing the reliability of water distribution systems for moderately large systems (10 - 50 nodes) with unreliable elements such as: pipe breaks, malfunctioning pumps, and out-of-service transmission and treatment facilities.

A comprehensive literature review of reliability methods and measures from a number of fields is presented. The objective of this review is to understand previously used measures and the methods used to calculate them, to integrate measures and methods from different fields, and to suggest other measures and problems to be addressed in a reliability analysis.

The water distribution is modeled as a flow-carrying network, with reliable supply and demand nodes and unreliable links. Link failures are assumed to be statistically independent. Analytical methods are identified and/or developed to assess the following measures on these networks: (1) the probability that all demand points in a system are connected to a source; (2) the probability that a given demand point in a system is connected to a source; and by assigning a capacity limit to each link in the system (3) the probability that a system can meet a specified level of flow at each demand point. Two sample systems are analyzed with these methods.

A stochastic simulation program is developed which calculates a number of reliability measures for networks with reliable supply and demand nodes, unreliable links, and water storage tanks of finite volume. Link failures are again assumed to be statistically independent. Different probability distributions are postulated for the distributions of (1) time until failure and (2) time until repair of individual elements. Three sample systems are analyzed with this program.

Finally, an overview of a general methodology for the reliability assessment of an existing water distribution is presented.

ACKNOWLEDGEMENTS

This project was funded by the National Science Foundation, by a grant concerning Stochastic Methods for Water Resources Management. *(NSF Grant CEE 8408369)*

We would like to extend our thanks to I. W. Dingwell, President of DIS/ADLPIPE, Inc., for allowing us the use the VAX upon which most of the work for this project was performed.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

1.0 INTRODUCTION

Most people, at least in developed countries, take it for granted that when they turn on their faucets, water will come out. Yet the reliability of water distribution systems is becoming of increasing concern to water system designers and operators because of the decreasing availability of public money and the increasing age of many systems. Reliability analyses of water distribution systems involve much more than just cataloging failure events. Also of interest are the occurrence of other water supply problems such as low pressure, the amounts of water not supplied, the spatial and temporal distribution of reduction or loss of service, and perhaps even the economic and other consequences of the lack of water distribution supply reliability.

Traditionally, water systems are designed according to certain guidelines which represent reliability considerations. For example, a common guideline is that each demand point must be supplied from at least two directions. Additional analyses may be performed to ensure that the system can operate under certain contingencies, *e.g.* with one pump inoperative all demand nodes must have 40 psi under conditions of average demand. These fixed criteria, however, treat probabilistic phenomena of reliability deterministically. Usually only a few contingencies can be examined in these analyses, and it may not be clear that the contingencies picked for analysis are indeed the ones of most concern. Contingency analysis is also unsatisfactory for choosing among proposals for reliability improvements since alternative designs for the same system may all meet the stated criteria.

This report is concerned with identifying, developing, and applying methods for calculating probabilistic reliability measures for water distribution systems. Our contention is that the reliability of a system involves stochastic events, so reliability can only be assessed by probabilistic (not deterministic) measures. Thus we focus on reliability measures such as the probability that a given demand point receives sufficient flow, or the probability that all demand nodes in a given system are connected to a source. In contrast to deterministic methods, probabilistic methods implicitly or explicitly account for the likelihood and effects of each system contingency. These measures also allow for easy comparison and ranking of similar systems on the basis of reliability.

This report will focus on the reliability of the distribution system only, not on external causes of unreliability such as insufficient supply or catastrophic system failure. Service losses will include pipe breaks, malfunctioning pumps, and out-of-service transmission and treatment facilities. Some of the methods explored will also include loss of service due to insufficient internal storage (*e.g.* water tanks). Events external to the distribution system which can also cause service losses, such as total power failure, drought, earthquakes, or enemy attacks, are not within the scope of this report.

Another objective in this report is to identify methods suitable for calculating reliability measures on moderately large systems - those of 10-25 demand and transmission nodes and 10-50 links. Thus more attention will be paid to fast methods for simple (and possibly approximate) models, than to very detailed methods which can be applied to systems with only a very few components.

Reliability methods in the literature fall into two broad categories, namely those for the assessment of the reliability of existing systems, and those for the design of new (reliable) systems. The focus of this report will be on methods for reliability assessment. Part of the rationale for this choice is that reliability problems are likely to occur in older systems, with demands exceeding their original design specifications and with the individual component reliabilities changing due to aging. Design rules which will ensure a maximally reliable network for a certain reliability measure on a certain class of network are of little help to a system operator with a system unlike this class of network. Our final audience for this work is water system operators and analysts with the need to explicitly identify the extent of problems they have, or fear they will soon have, in an existing system. Also, approximate methods of reliability assessment are a necessary first step for the design of reliability-optimal systems.

In summary the objective of this report is to present:

* methods to assess the reliability of water distribution systems,
* using probabilistic measures,
* for moderately large systems,
* with unreliable system components.

A comprehensive literature review on the subject of reliability of water distribution systems is presented in Chapter 2. This review is organized around the many different measures of reliability employed, both in the field of water system design and operation and in other fields such as Operations Research, Communications, Safety Engineering, and Electrical Engineering. Special attention is paid to methods from other fields which may be applicable to water distribution systems.

The definition of reliability measures is a crucial issue in the assessment process. Methods exist to calculate a number of reliability measures, but these measures may not be sufficient to characterize the reliability of the system in question. On the other hand, many interesting measures can be defined which can not be calculated, or which can not be calculated without a great deal of time and expense. Throughout this report an attempt is made to "walk the fine line" between these two extremes, that is, to identify useful measures which can be computed with reasonable effort.

In Chapter 3, analytical methods are identified and/or developed for assessing the following measures:

* the probability that all demand points in a system are connected to a source,
* the probability that a given demand point in a system is connected to a source, and
* the probability that a system can meet a specified level of flow at each demand point.

The water distribution system is modeled as a flow-carrying network, allowing the use of network analysis methods and the calculation of these measures for moderately large-sized systems. Following the discussion of these analytical methods is a discussion of stochastic simulation methods in Chapter 4. Simulation methods, although generally more complicated and requiring more computer resources than analytical methods, allow greater flexibility in the detail of the system model and in the measures of reliability employed. Both Chapters 3 and 4 contain the analysis of sample water distribution systems with the methods discussed. In conclusion, Chapter 5 presents an overview of a methodology for using the previously identified techniques to perform a reliability analysis of an existing water distribution system. Suggestions for further work on reliability assessment of water distribution systems are also included in this conclusion.

2.0 DISCUSSION AND LITERATURE REVIEW OF RELIABILITY MEASURES AND METHODS

2.1 INTRODUCTION

Reliability is a word used in many fields. Reliability calculations have been performed for systems ranging from a 2 cent fuse to nuclear power plants and the Apollo space missions. However, like many words with both common and technical uses, the meaning of "reliability" can change from application to application. The statement "my spouse is reliable" is understandable in general conversation, however the statement "my spouse is 95.2% reliable" is so nonsensical it conveys little information. Unfortunately, the statement "this water distribution system is 95.2% reliable" conveys little more information, without knowing in detail how the reliability was measured, from what data, by what method, and with what model.

A reliable system is one which performs consistently as expected or required; in this sense reliability is a system objective. To measure the extent or degree to which the system does perform consistently, reliability performance measures are needed. For any particular project the choice of measure or the design of a new measure is, however, not always a a straightforward process.

The first consideration for measuring reliability involves defining when the system is performing and when it isn't. Desired performance can vary both spatially and temporally. Perhaps it does not matter if there is a 10 minute period of no water supplied if it occurs at 3:00AM, but a 10 minute period of water loss at noon may be extremely annoying to the customers. Water loss may also be of more concern when it occurs randomly, than when planned for in advance. Additionally intermittent loss of water for short periods may not be of great concern in the parts of the system supplying mostly residential users, where even short unplanned outages to the part of the system providing cooling water to a refinery may be disastrous. System performance measures may also need to account for the occurrence of reduced service periods, when water is still supplied but at a flow rate or pressure less than normal and/or desired by the users. Thus it is desirable for the reliability performance measures chosen for an application account for the important spatial, temporal, and use variations in the system.

On the other hand, if the reliability measures chosen are too complex they may be impossible to calculate. Additionally, reliability is a probabilistic phenomena. It is simply not possible to pinpoint where every pipe break within the next year is going to occur. Thus some aggregation of events, over space and/or time, must be done for these measures to have any meaning. As stated in the introduction, we must "walk the fine line" between simple measures which may be easily calculated and more complex measures which better reflect the expected performance and variation of the system but are more difficult to calculate.

The purpose of this chapter is to identify reliability performance measures (also called reliability definitions or reliability indices) that have been used for the analysis of water distribution systems. Additionally, applicable measures from other fields, such as Operations Research, Electrical Engineering and Communications, Electrical Power Systems Engineering, and Nuclear Engineering, will be discussed. Available methods for the calculation of these measures will also be summarized.

Focusing on reliability measures serves a number of purposes. First, the initial step in using any published method is to determine the measure being calculated. Emphasizing reliability measures provides a way to understand the reliability literature. Second, a careful examination of methods used in different fields reveals that some seemingly dissimilar methods are, in fact, calculating the same reliability performance measure.

Taking an example from Electrical Engineering, a number of sophisticated techniques have been developed to calculate reliability for "multi-state systems". Multi-state systems are those which have intermediate states between failure and full operation. In Electrical Power Systems Engineering, several methods exist for calculating the reliability of electrical power generating systems (or grids), composed of a number of connected generating plants. Upon reflection it becomes clear that in terms of generating capacity, an electric grid is also a multi-state system, although the term as such is not used in the field of Electrical Power Engineering. Examining reliability measures allows the integration of methods from a number of different fields.

Third, looking at a number of methods from differing contexts highlights the range of issues that may be involved in a reliability analysis. Thus even for practitioners more interested in applying the methods discussed in the later chapters than in the background of these methods, a fast reading of this chapter may suggest other measures that should be used for their particular problem or other concerns to be addressed in the analysis.

One last note of caution however. The focus on measures and methods previously used for reliability analyses is in no way meant to imply that these measures are all that can or should be developed. The temptation must be avoided to accept what we can compute as what we should compute. As computational power increases, from both increased computer power and better algorithms, more complicated and less aggregated measures should be developed and employed. Equally crucial to a good reliability analysis, is remembering that important reliability issues, even if not easily quantified, can not be ignored. Currently computable reliability measures must not become a "way of life" based on our ability to compute instead of on our satisfaction with these measures.

2.2 SYSTEM AGGREGATION AND RELIABILITY MEASURE CALCULATIONS

As mentioned in the last section, reliability measures must aggregate events over space and/or time for these measures to have any meaning. The level of aggregation of a particular measure also provides a useful way of categorizing these measures and the methods for their computation.

In general, the more highly aggregated the system, the simpler the analysis and the more detailed the information obtainable. For example, for a highly aggregated system with one source directly connected to a few demand nodes it is often possible to calculate the full probability distribution of some reliability measures. For a detailed water distribution network however, it may only be possible to calculate the mean of this measure. Three commonly used levels of aggregation are:

(1) lumped supply, lumped demand,
(2) delineated supply and transmission, lumped demand, and
(3) both supply and demand delineated (networks).

Even for most network measures, state-of-the-art methods usually deal with a maximum of 10 to 100 nodes. Thus for most applications individual water consumers must still be aggregated into a moderate number of "demand nodes".

Within each category a number of different performance measures have been defined. Generally, for a given level of aggregation the applicable measures differ in the detail to which the processes of failure and repair of individual components are modeled. Again, in general, more assumptions imply easier computations, but less detailed reliability information.

The following sections discuss reliability methods applicable to water distribution systems, for each of the three categories of aggregation. Each section is organized around reliability performance measures, in order of increasing complexity. Methods for the calculation of each performance measure are sketched, and sources for more detailed explanations are reviewed. Guided by the reasons for focusing on reliability measures, each section is intended to help a reader *understand* available measures and methods at this level of aggregation, *integrate* the methods from different fields, and *suggest* further measures and concerns to be addressed in a reliability analysis of a specific system.

2.3 SYSTEMS WITH LUMPED SUPPLY, LUMPED DEMAND

For a general investigation of unreliability, due only to lack of sufficient system water supply, the water distribution system can be modeled as an area of "total supply" connected to "total demand" (Figure 2-1). Models of this sort have been widely used in the field of Electric Power Engineering to determine the required amount of system generating capacity needed to meet electrical demand in an area.

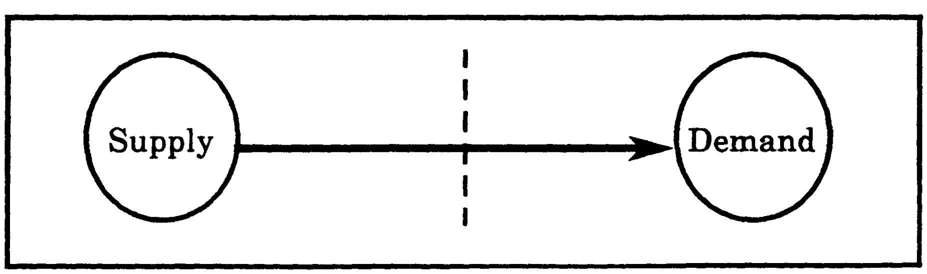


Figure 2-1 System with Lumped Supply, Lumped Demand

Electrical power generating systems, sometimes called electric power grids, consist of a number of connected electrical generating units. These units may be modeled as operating (on-line) or not operating (off-line). Some units may also have intermediate states, namely they may be able to generate power at number of levels. Outage of individual units may be planned or random occurrences. Random failure of individual generating units is usually assumed to occur independently. Thus over time the system will exhibit a range of generating capacities. Depending on the demand occurring during each period, the system will or will not have sufficient generating capacity in that period.

Reliability Evaluation of Power Systems, by Billinton and Allan [1984], contains a comprehensive review of a number of methods for the calculation of reliability performance measures for power generating systems. Relevant measures include:

* loss of load expectation (LOLE) - the expected number of days in a specified period on which the load will exceed the available capacity,
* loss of energy expectation (LOEE) - the expected amount of energy in a specified period demanded but not supplied,
* frequency of load (or energy) loss - the expected number of occurrences per unit time of a load (or energy) loss,
* expected cycle time - expected duration between occurrences of a given capacity, and
* average duration of a given capacity.

These measures can be readily generalized to water systems by substituting *water demand* for *electrical load,* and the total *volume of water* supplied for total *amount of energy* supplied.

For LOLE and LOEE each generating unit in the system is assigned one or a number of generating capacities, and a probability of operation at each capacity. The demand on the system (load) is assumed to vary according to some probability distribution. The methods proceed by convolving the capacity outage probability distribution with the load duration distribution.

The capacity outage distributions can be compiled by fairly straightforward methods. For every possible combination of fully operating, partially operating, and failed units, the capacity of this combination and the probability of its occurrence is calculated. Summing up the probabilities for all combinations with equal generating capacity, the probability of occurrence for each possible capacity level is found. From this distribution, the *cumulative* probability distribution of capacity (the probability of having a capacity equal to or greater than a given capacity) is also easily calculated (Figure 2-2). For systems with many generating units, recursive and transform methods exist for the calculation of these distributions.

System demand distributions can be constructed from a variety of data, depending upon the desired analysis. A *daily peak load curve* is obtained by arranging the observed daily peak load values for a period of time in descending order, and calculating the cumulative percentage of loads that exceed given values. A *load duration curve* is similarly obtained using hourly load data. (Figure 2-2)

For a given available capacity the percent of time the demand exceeds this capacity can be taken directly from the load duration curve (demand distribution). Thus by convolving the two curves, the expected loss of load can be calculated. For a given available capacity the total amount of energy not supplied is given by integrating the demand distribution between the actual demand satisfied and the demand required. Thus the loss of energy expectation can also be found from these two curves. These LOLE and LOEE methods have great

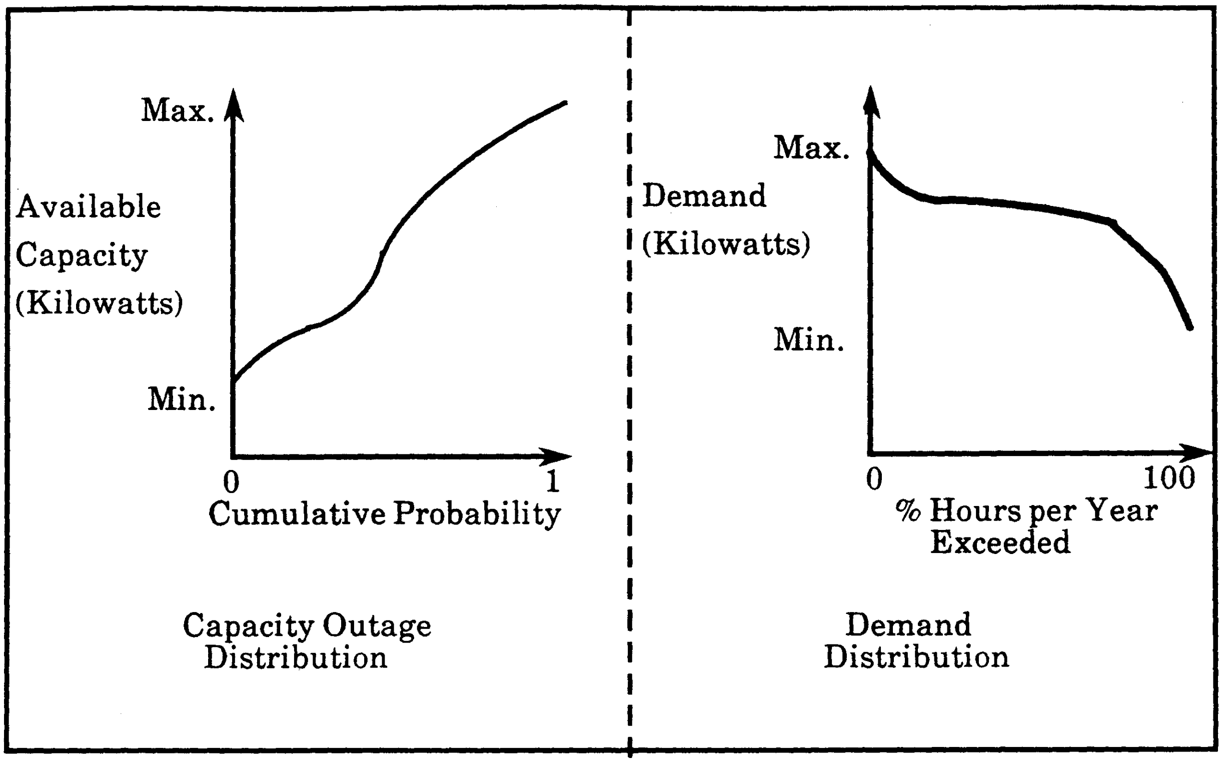


Figure 2-2 Example Capacity Outage and Load Demand Curves

flexibility and can also be extended to account for outages due to planned down time for maintenance, power storage, outages due to limited energy to produce the electricity, and other complex circumstances.

Hobbs [1985a] presents a LOLE reliability analysis of water system capacity for a small system that can be easily collapsed into areas of lumped supply and lumped demand. Hobbs also continues the analysis to calculate the frequency and average duration of the failures of this system. Shamir and Howard [1981] also present similar methods developed specifically for water distribution systems. They define three reliability factors (*i.e.* measures) as follows:

* discharge reliability factor: RC = 1 - (lost capacity/total required capacity)^n (n a user-supplied constant),
* volume reliability factor: RV = 1 - (lost volume/total required volume), and
* overall reliability factor: RF = (RC + RV) / 2.

The discharge reliability factor (for n = 1) is the random variable used to calculate the expected loss of load above, and the volume reliability factor is the random variable used to calculate the expected loss of energy. Howard and Shamir however, determine the entire probability distributions for these measures, using derived distribution methods from applied probability. In their analysis, demand is assumed constant and loss of capacity is assumed to be due only to pump failure. Howard and Shamir also consider the factors of storage, pumping capacity, and repair times on the reliability distributions for a lumped system. In a later paper, Shamir and Howard [1985] use a method similar to LOLE to determine the expected reliability (defined as the ratio of supply over demand) over the next 50 years for Seattle, Washington.

Reliability measures of capacity frequencies, durations, and cycle times are calculated by methods called, logically enough, frequency and duration (or F & D) methods. These methods rely on the theory of discrete state, continuous time Markov processes. For these models, a "state" is defined for each distinct combination of operational generating units. The time spent in each state i is usually assumed to be independent of all other states and events, and exponentially distributed with transition rate λ_i. For an exponential process, the reciprocal of the transition rate for a state gives the mean occupancy time of that state (mean time spent in a visit to the state). Transition rates are usually modeled as remaining constant over time. Solutions to these Markov models provide values for capacity frequencies, durations, and cycle times. Markov models are usually applied to less aggregated systems, so some further discussion of this topic and related references will be saved for the next section.

These highly aggregated models are probably of more use in power system capacity reliability studies than in water system reliability analyses. Because of the facts that (1) most local electrical systems are highly interconnected, (2) electric transmission line losses (both of power and voltage) are often negligible over moderate distances, and (3) electricity travels at the speed of light, to a good approximation supply anywhere in the system can reach demand anywhere else in the system. Thus aggregate analyses are useful since the exact physical location of the generating plants and the electric consumers is of little concern. However for water systems, (1) interconnections are relatively sparse, (2) significant volume and pressure losses do occur along water pipelines, and (3) water in pipes does not travel at the speed of light, thus the spatial distribution of a water system is usually of prime importance to any reliability analysis. However, these methods for aggregated systems could be of interest to water system analysts for a first analysis of water system reliability or as part of a larger analysis.

Other, more complicated, reliability measures could be developed for water distribution systems with lumped supply and demand, particularly by focusing on measures less aggregated in time. However, for most analyses it would more important to look at models less aggregated over space, as in the following sections.

Further information about these methods is available in an earlier book by Billinton and Allan, not restricted to power system reliability, called *Reliability Evaluation of Engineering Systems* [1983]. Another often quoted comprehensive text in this area is *Reliability Modeling in Electric Power Systems* by J. Endrenyi [1978]. Additionally, as a guide to the journal literature two comprehensive bibliographies concerning probability methods in power engineering, compiled in 1972 and 1978, can be found in the IEEE Transactions, Power Apparatus and Systems (PAS) [Billinton, 1972; IEEE Subcommittee on the Application of Probability Methods, 1978].

2.4 SYSTEMS WITH DELINEATED SUPPLY AND TRANSMISSION, LUMPED DEMAND

One step in analyzing a water distribution system in more detail than in the previous section is to focus on the bulk transmission system. Typically, such studies involve models with separate elements for each water source (*e.g.* reservoir, aquifer, well), treatment plant, connecting pipeline, and only one large demand area. Figure 2-3 shows such a model, taken from the water supply system of Dordrecht, The Netherlands, and analyzed by Tangena and Koster [1983].

Many reliability methods can be (perhaps loosely) categorized as intended for systems at this level of aggregation. At this level of aggregation it is still possible to use analytical methods, either exact or approximate. Several studies of water systems (described later in this section) have been done for such systems, mostly using derived distribution methods from applied probability. Many of the methods from Electrical and Systems Engineering can also be classified into this category, since an electrical component can be thought of as carrying electricity from one point ("the source") to another ("the demand area") through a delineated sub-circuit ("the transmission system"). Analyses of electric power generating systems for which all elements have not been aggregated into one large supply node and one large demand node (as in the previous section) have also been performed. Methods for these analyses of electrical power systems are often called transmission reliability methods.

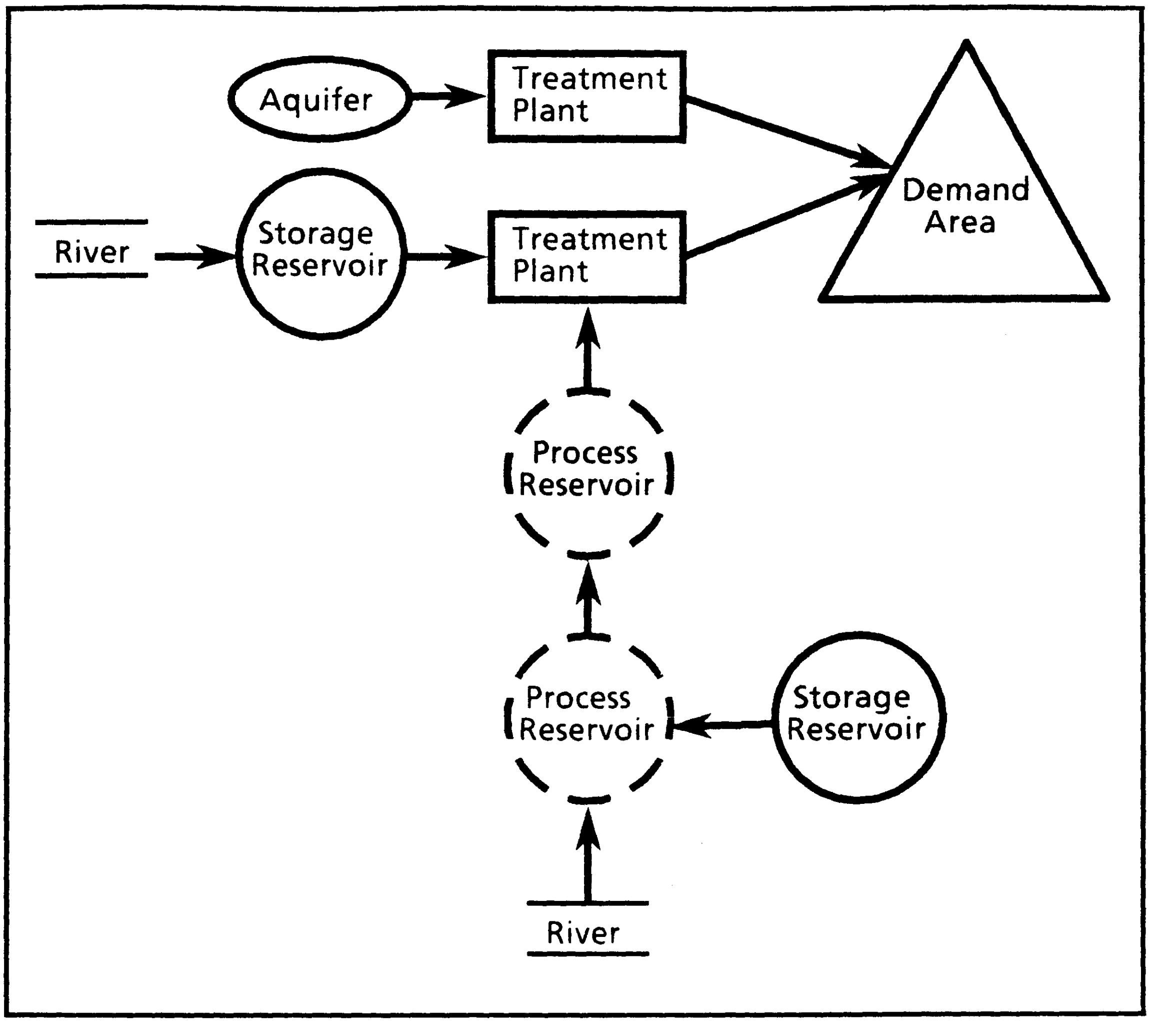


Figure 2-3 Bulk Transmission System
(taken from Tangena and Koster [1983])

At this level of modeling each component is usually described as either operational or failed. The state of each component is represented by a binary random variable x_i such that:

$x_i =$ 0 if the component is failed,
1 if the component is operational.

Depending upon which components are operational and which are failed, the system is then classified as operational or failed. Then the state of the system can also be represented by a binary random variable y, which is a function of the states of the components. Again:

$$y = \begin{array}{l} 0 \text{ if the system is failed,} \\ 1 \text{ if the system is operational.} \end{array}$$

Most of the methods for reliability analysis of this type require the systems to be coherent, that is the failure of an additional component can not cause a failed system to become operational.

Within this model, there is considerable flexibility in the definition of an operational system. For example, a system may be operational if the full demand of the demand area is met, if 90% of the maximum demand is met, or if the demand area is still connected to some source (disregarding the amount of water or electricity supplied). As long as coherence, as defined above, is maintained the analyst may chose any definition of an operating system. The choice should be based on the needs of the problem, and on the ease to which a given combination of failed and operating elements can, according to the definition, be identified as providing an operating or failed system. Since the methods to be described in this section will work for almost any definition of an operational system, the generic term of an "operational system" will be used without further clarification.

The reliability of this system is measured as:

* the probability that the system is operational (Pr [$y = 1$]).

Many of the reliability analysis methods for this model were developed originally for electrical circuits. In much of the Electrical Engineering literature this definition of reliability as the probability that a system operates is used implicitly. (For a nice discussion of the history of this field see Barlow [1984].) A summary of reliability methods in Electrical Engineering up until the mid-seventies and the starting point for almost all work in this area since then is the text "Statistical Theory of Reliability and Life Testing: Probability Models" by Barlow and Proschan [1975]. Most of these methods are directly applicable to water distribution systems at this level of aggregation. Additionally, for bulk water supply systems Hobbs [1985b] provides a qualitative discussion of reliability models and measures.

Methods for reliability, as defined above, vary in how the individual component reliabilities are described. At the simplest level, the probability that each component operates is given as a constant p_i, with the probability of failure $q_i = 1 - p_i$. These values may represent, for example, the percentage of time over a long period a repairable component operates, or they may represent the probability a component will last for a specified period of time. An additional assumption is usually made that individual components fail independently. Thus reliability measures based on a system with component reliabilities as just described are highly aggregated over time. Methods for calculating the probability that the system is operational include:

* enumeration of all possible combinations of component states,
* recognition of systems composed of special structure, such as systems connected with only series and parallel connections (see Figure 2-4), and
* decomposition.

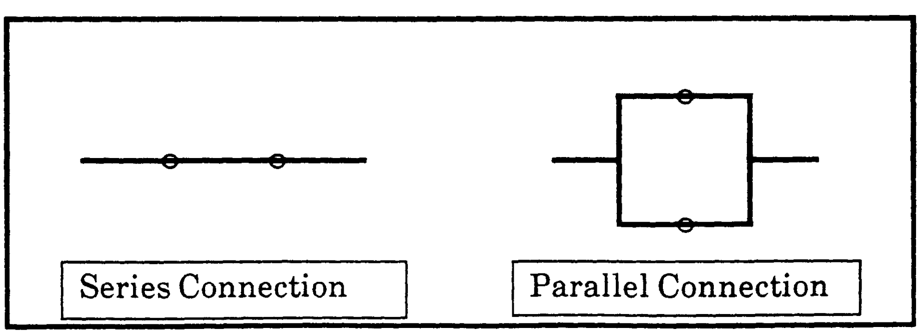


Figure 2-4 Series and Parallel Connections

Decomposition is an application of conditional probability. For example, let:

a be a system component,
p_a be the probability of operation of component a,
q_a be the probability of failure of component a,
y be the probability of operation of the system (or system reliability),
$y(a)$ be the system reliability with a operational, and
$y(\underline{a})$ be the system reliability with a failed.

Then $y = y(a) \cdot p_a + y(\underline{a}) \cdot q_a$. If the system can be decomposed into systems with simple or special structure, decomposition can save considerable computational effort.

Many methods, based on the above model, have been developed to design "optimally" reliable systems. Tillman, Hwang, and Kuo in "*Optimization of Systems Reliability*" [1980] present a review of these methods. In Chapter 2, they identify three reliability optimization problems, namely:

(1) allocating reliability among the components of a system so as to maximize the overall system reliability (with or without cost constraints),

(2) allocating redundancy to maximize system reliability (usually by adding components in parallel), and

(3) minimizing system cost subject to reliability constraints.

More detailed but more complicated models can be developed by representing the time until failure of a component as a probability distribution. These "lifetime" distributions can then be used to obtain values for the p_i's above, as the probabilities that the individual components will last for a specified period of time. Then in addition to the overall probability that the system will operate, time-related reliability measures can be calculated such as:

* mean time to (system) failure (MTTF),
* variance of system failure times, or even
* system lifetime probability distribution.

For many systems, components can be repaired after they have failed. For such systems both the time until failure and the time until repair can be treated as random variables. Systems can also be inoperative due to planned maintenance, and these outages also can be included in reliability measures. In the context of "repairable and maintainable systems" measures of the reliability of the system are often called availability measures. Such measures include:

* mean time between failures (MTBF) or cycle time,
* mean time to repair (MTTR),
* mean time between maintenance (MTBM),
* instantaneous availability (Pr [system operates at random time t]),
* time-averaged availability,
* steady state availability,
* availability for an interval (0,T),
* frequency of failure events - the expected number of failures per unit time.

Although methods from different fields for calculating availability measures may employ very different terminologies, most of the methods are based on similar underlying concepts. As touched upon in the last section, these methods rely on the theory of discrete state, continuous time Markov processes. At each point in time the system can be in one of two discrete states: operational or failed. It is usually assumed the time spent in each state is independent of all other states and events, and exponentially distributed. The transition rate from operational to failed is usually denoted as λ, and the rate from failed to operational as υ. Figure 2-5 presents the transition rate diagram for this continuous time Markov process. Again, for an exponential process, the reciprocal of the transition rate for a state gives the mean occupancy time of that state (mean time spent in a visit to the state). Thus the system spends on average $1/\lambda$ time units before breaking, and $1/\upsilon$ time units in repair. Because the theory of such two state continuous time Markov processes has already been done, and many of the formulas for the above defined reliability measures tabulated, the major effort for many such analyses is the calculation of the failure and repair time distributions.

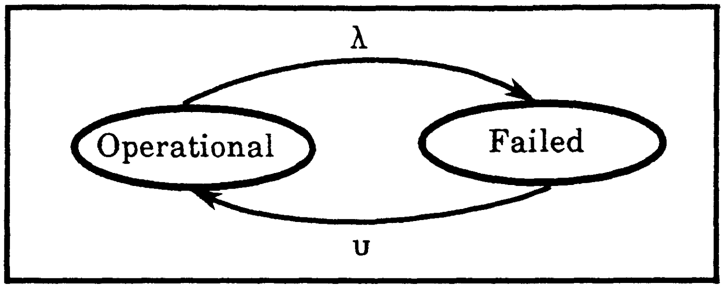


Figure 2-5 Markov Model for System Availability

For simple failure and repair distributions many availability measures may also be calculated from simple derived distribution methods from applied probability. Tangena and Koster [1983] illustrate this approach for the reliability analysis of the system in Figure 2-3. In this analysis, both the failure and the repair distribution times are assumed to be distributed exponentially. Some of the frequency and duration (F & D) measures discussed in Billinton and Allan [1983 and 1984] are also based on this approach. Additional methods for more complicated distributions, as well as for more complicated dependences going beyond traditional Markov model systems, are also discussed in Billinton and Allan as well as in the text by Barlow and Proschan [1975], and a bibliography by Lie, Hwang, and Tillman [1973].

Another step can be taken to increase the detail of these analyses, by allowing the systems to have numerous operational states. For water and electric power systems these states would usually represent current system capacity in terms of amount water (or electricity) supplied per unit time. Many of the previously defined reliability and availability measures can be generalized to the multi-state case, to define measures such as:

* probability that the system is in state i (supplies capacity i),
* frequency of occurrence of state i, and
* cycle time between occurrences of state i.

In addition, capacity related measures can be calculated such as:

* volume shortfall - total amount of water (energy) demanded but not supplied,
* time-averaged capacity shortfall - average level of capacity demanded but not supplied,
* Pr [demand exceeds supply in interval (0,T)],
* Pr [demand exceeds supply at any time t], and
* Pr [demand is exceeded for more time than a given value t].

Methods for calculating these multi-state measures are discussed in the previously mentioned sources. Shamir and Howard [1985] use derived distribution methods, to calculate the complete probability distribution of the volume shortfall for a small bulk water transmission system.

There are a number of more complicated reliability measures that could be developed for systems at this level of aggregation. One obvious extension is to tie these reliability measures to economic consequences. At this level it should then be possible to develop analytical methods for calculating such quantities as expected cost of water losses. Such methods could then be used to perform cost/benefit analyses on possible improvements to the system. Also availability calculations which account for planned outages could be used to develop optimal maintenance schedules for the elements in the water transmission system.

Currently available methods for systems with lumped demand are have been shown by several papers mentioned above to be useful for water distribution system reliability analysis. At this level of aggregation a number of fairly complex measures can still be calculated analytically. Many of these methods could be extended to systems with 2 or 3 demand areas with minimal difficulty. However for a system with fully delineated demand, other methods become necessary.

2.5 SYSTEMS WITH DELINEATED SUPPLY AND DEMAND – NETWORKS

A water distribution system can usually be described in detail as a network of supply and demand points connected by pipes and pumps. Borrowing from graph theory, the supply and demand points are frequently referred to as nodes, and the connections as arcs or links. The water distribution network below (Figure 2-6), which is analyzed extensively in later chapters, is an example of a model of this type.

The reliability of the individual components of the network are described in many of the same ways for full network systems as for more aggregated systems, leading to many similarities between network-oriented reliability methods and the methods discussed in the previous section. However, more variations are possible for the definitions employed for when the system operates. For example, one definition of an operating system is again when demand at every node is met. However, for a network model an system may be described as operational if demand for only one specified node is met, or if demand for some specified subset of nodes is met, regardless of the states of the other nodes. In the previous section, where the one demand node was the only demand node such finely tuned definitions were not possible.

Network reliability has been studied mainly in the fields of Communications Engineering and Operations Research. Much of the work has concerned the problem of network connectivity, for independently unreliable links, described

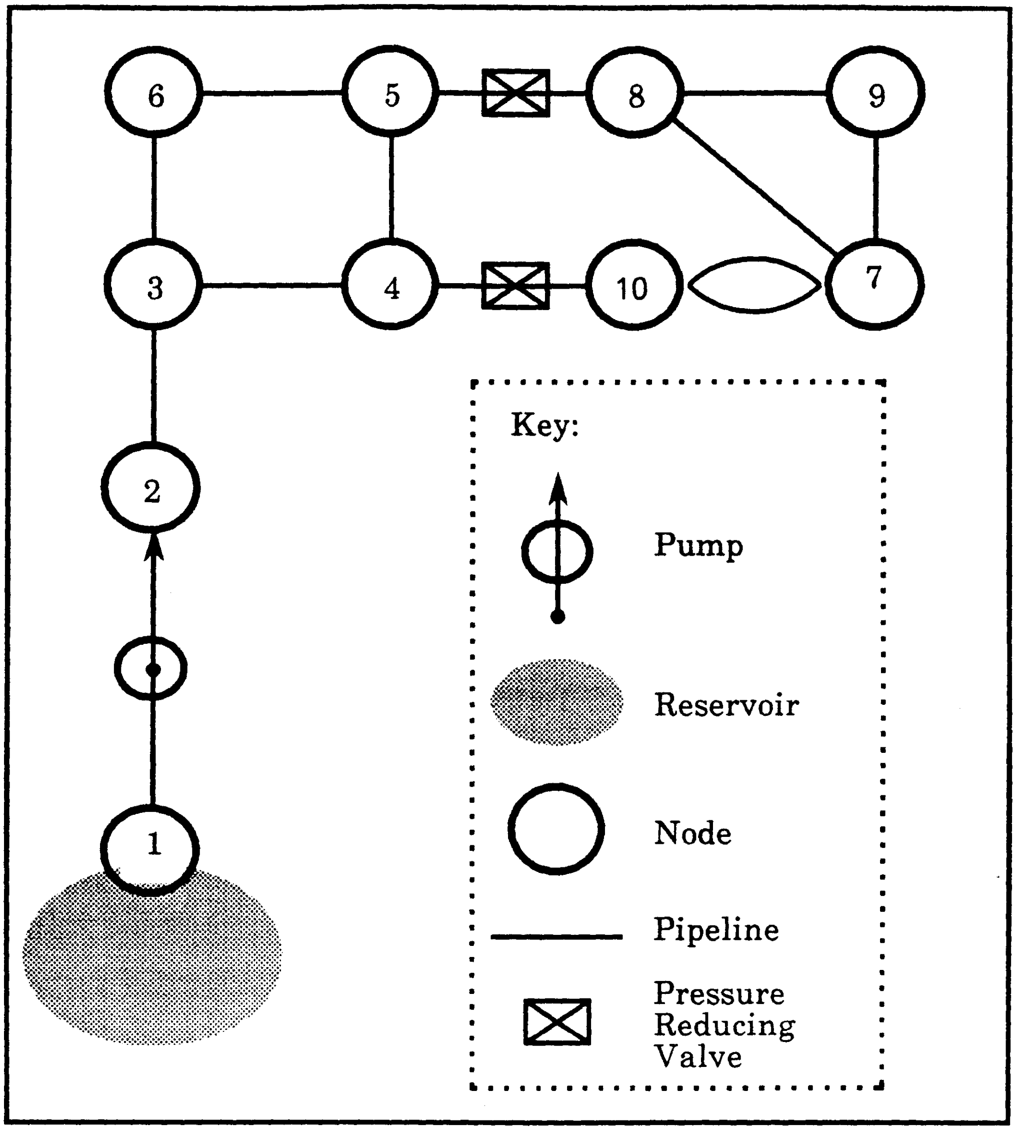


Figure 2-6 Water Distribution Network A

with p_i's and q_i's. In a paper on the complexity of network reliability computations, Ball [1980] identifies and classifies a number of reliability measures previously used in the literature. These measures include:

* probability that two specified nodes (s and t) in a network are connected,
* probability that a specified subset of nodes in a network are connected, and
* probability that all nodes in a network are connected.

In addition, it is possible to allow the nodes as well to be unreliable. In water systems unreliable nodes could be used, for example, to model water tanks which sometimes run dry. Then the above measures can be extended to include the probability that specified nodes are connected *and* operative.

Despite the seeming simplicity of this model, Valiant [1979], Ball [1980] and Provan and Ball [1983a] have proved that the computations of the above straightforward definitions of reliability are NP-hard for general networks. NP-hard means that these problems can be shown to be at least as hard as a large class of problems for which no "easy" algorithms are known. (An "easy" algorithm in this context is used to mean that the number of computations required by the algorithm grows only polynomially with problem size.) Thus it is likely no polynomial-time algorithms to calculate these reliability measures for a general network can be found, implying exact analytical reliability methods for extremely large networks will always be cumbersome. However polynomial-time methods for networks with special structure, such as all series and parallel connections, have been developed. A discussion of methods for calculating these connectedness reliability measures follows.

Satyanarayana and Wood [1982] present a polynomial-time algorithm for the calculation of the above reliability measures on undirected series-parallel networks. Agrawal and Satyanarayana [1984] followed by presenting a polynomial-time algorithm for calculating these measures on directed series-parallel networks. These methods involve identifying structures, such a triangles or series connections, in the networks. The methods provide ways to reduce these structures to simpler structures, without changing the reliability of the system. Both of these algorithms allow series-parallel networks to be reduced to a single edge, with edge reliability equal to the reliability of the entire system.

Satyanarayana and Wood's algorithm is employed in Chapter 3, where it is described in detail.

Agrawal and Barlow [1984] present a brief but clear survey of existing concepts and methods for the calculation of these reliability measures from the perspective of Operations Research. They identify three methods that may be used, instead of complete enumeration, to calculate these reliability measures, namely:

* inclusion-exclusion,
* sum of disjoint products, and
* decomposition.

Inclusion-exclusion methods deal with minimal path sets, which are sets of links such that the failure of any one link in the set will cause some node in the specified subset to become disconnected. For any one path set i, the probability of the event that all links in the path set are operating [$Pr(A_i)$] can be found as the product of the p_i's for every link in the path set. These methods involve identifying each minimal path set, calculating $Pr(A_i)$, summing up the $Pr(A_i)$'s for each path set, and subtracting some terms to account for the fact that some links occur in more than one path set and have thus been double counted. Unfortunately, there are an exponential number of these correction terms.

Sum of disjoint product methods also involve identifying each minimal path set. Let p indicate the number of path sets. Let A_i indicate the event: all elements in path set i operate, and $\underline{A}_i$ the (complex) event: any element in path set i is failed. The probability that the network is connected is given by:

$$Pr[y=1] = Pr(A_1) + Pr(\underline{A}_1 A_2) + \ldots + Pr(\underline{A}_1 \underline{A}_2 \ldots \underline{A}_{p-1} A_p).$$

In this method there are only p terms, however the work involved to calculate each term is usually exponential.

Decomposition (also called factoring) was described in the previous section. A general network can decompose into an exponential number of terms, so again these methods are, in the worst case, exponential.

Agrawal and Barlow [1984] reference a number of algorithms for these methods. Other recent algorithms include: Rosenthal [1977], Buzacott [1980], Johnson [1984] and Provan and Ball [1984]. Ball and Provan [1983b] present approximate methods for these network based measures. Some work on calculating these measures without independent failures has been done (see Zemel [1982] and references). Although most of these algorithms will work with any general network, most methods exploit some special network structure. Considerable computational saving may be obtained by chosing a method suitable to the network at hand. In Chapter 3, Rosenthal's algorithm of was used and is explained in more detail.

Reliability measures involving link capacities and limits on supply can also be defined for networks, although they have received somewhat less attention in the literature. Applicable measures include:

* probability that demand can be met at every node in the system,
* probability that demand can be met at some specified subset of nodes in the system,
* probability that demand can be met at any time t,
* probability that demand can be met within an interval (0,T),
* expected total shortfall in the system, and
* expected shortfall at any specified node.

Most of the methods for calculating these measures involve specifying a capacity or flow limit on each link in the network. The links are again assumed to be independently unreliable, and the individual reliabilities of each link are described by p_i's and q_i's. Some of the above connectivity type methods can be extended to this case, specifically Rosenthal [1977]. Later, Rosenthal [1981] developed a faster method for calculating these supply related measures on series-parallel networks. Other methods for measures of this type are presented by Aneja and Nair [1980], Willie [1979], and Lee [1980]. Shogan [1982] presents a method for these calculations when supply and demand are also random. Carey and Hendrickson [1984] present approximate methods for these measures. In Chapter 3, Lee's algorithm is adapted and applied to water distribution systems modeled as capacitated networks. Details of Lee's algorithm are presented there.

In addition to the probabilistic measures described above, a number of deterministic reliability measures, dependent only on the layout of the network, have been defined. These include such measures as:

* minimum cut-node set - the smallest set of nodes such that if any one of these nodes are removed from the network, some node becomes disconnected,
* minimum cut-link set - the smallest set of links such that if any one of these links are removed from the network, some node becomes disconnected, and
* diameter - the maximum number of links in simple (non-cyclical) paths between any two nodes in the network.

These deterministic measures are often referred to as measures of vulnerability. Deterministic measures are most useful for characterizing the reliability of networks, for a case where the disruptions are caused by an agent with some intelligence and information about the network, *e.g.* for networks under enemy attack. However, the operations of finding cut-node and cut-link sets are parts of many probabilistic reliability methods. Additionally, some of these measures are useful for examining the reliability of networks under catastrophic conditions, for example as a step in the analysis of the likelihood that a water distribution network is able to survive an earthquake. Wilkov [1972] provides a comprehensive discussion of a number of deterministic reliability measures used for reliability analysis of computer networks. Wilkov discusses some of the probabilistic measures previously discussed in this section as well.

Lastly, another set of network-oriented reliability methods is fault tree analysis, frequently used for safety analysis in the fields of Nuclear and Chemical Engineering. A fault tree is a description, as a network, of the ways a structure (*e.g.* nuclear reactor) can fail. Thus, in contrast to the reliability methods discussed above, a major part of fault tree methodology is concerned with generating a complete list of the nuts and bolts events which could cause the structure to fail. For example, Figure 2-7 shows an example of a fault tree for the collapse of a bridge.

When probabilities are assigned to each element in the fault-tree, the probability of each failure (branch) occuring, and thus the total probability of failure of the structure can be calculated. If a number of types of failure are possible for a structure, each with different consequences, an event tree (similar to a fault-tree) may be constructed. An event tree diagrams all possible paths following from a top event, through a number of possible subsequent events, to the possible consequences.

Analysis of fault and event trees also involves the identification of cut-node sets in the trees, and the summation of the possible failure modes. Probabilities of failure of the system within a specified period of time, and mean time to failure are frequently used measures for fault and event tree analysis. Willie [1978] has shown that the analysis of fault trees, without NOT-gates, is equivalent to probabilistic network analysis (as described at the beginning of this section). However, since fault trees for actual systems are very large (thousands of nodes)

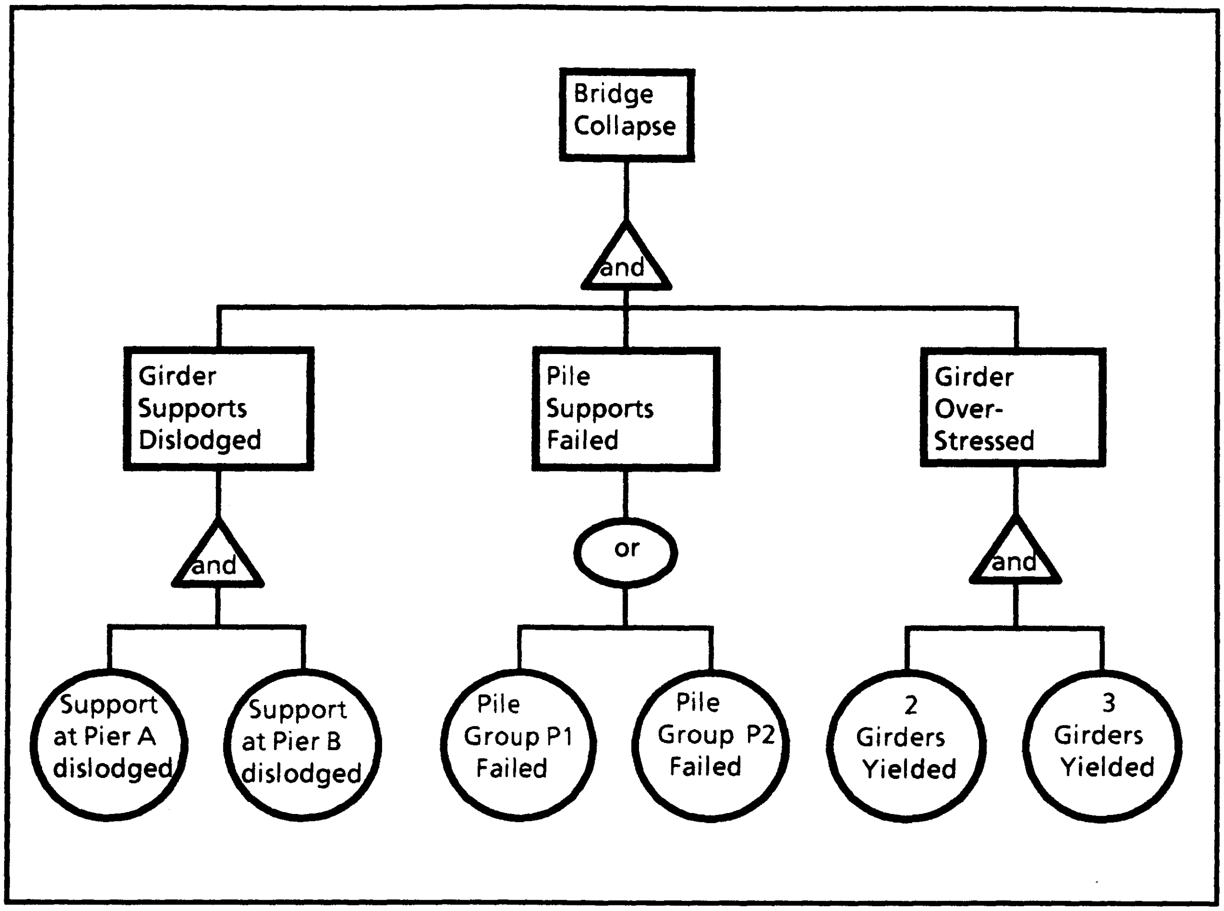


Figure 2-7 Fault Tree For Collapse of Bridge

(Adapted from Ang and Tang [1984] pp 487)

reliability measures for most fault tree models can only be approximated. For example, simplifying assumptions are frequently made to enable a lower bound on the system reliability to be found. In addition, many of the methods in this field involve heuristic rules such as disregarding failure events with minimal effects or extremely low probabilities.

A large literature on fault and event tree modeling and analysis has been developed. Two useful introductions to this field are *Reliability Engineering and Risk Assessment* by Henley and Kumanmoto [1981] and *Probability Concepts in Engineering Planning and Design* (Volume 2. Chapters 6 and 7) by Ang and Tang

[1984]. A paper by De Jong, Snijders, and Wedmijer [1983 in Dutch], is reported by Tangena and Koster to use failure-mode-and-effect analysis a part of a methodology applied to analyze the reliability of the drinking water supply of the city of Rotterdam.

Most water distribution networks involve several water sources, transmission elements, and a large number of demand points, thus network based reliability measures are potentially very useful for analyzing the reliability of these systems. Chapter 3 focuses on three network based reliability measures which can be calculated analytically.

In the future, the analytical methods for the calculating measures that are less aggregated temporally should be developed for networks. Network reliability measures, such as the frequency and duration measures discussed in the previous section, are desirable. The simulation presented in Chapter 4 presents a non-analytical way of determining the values of reliability measures such as mean time between failure, and mean time to repair.

3.0 ANALYTICAL METHODS

3.1 USE OF ANALYTICAL METHODS FOR RELIABILITY CALCULATIONS

Practical analytical methods for calculating reliability measures defined on water distribution networks can be adapted from network reliability methods developed in other fields. The aim of this chapter is to identify several reliability measures that are both useful in assessing the reliability of a water distribution system, and can be calculated via methods requiring reasonable amounts of input data, computer resources, and labor to use. This chapter will focus on network based methods, because it is only these methods that allow reliability calculations on networks of more than a few nodes. Analytically calculated measures should allow fast initial assessments of the water system and easy analysis of the sensitivity of these measures to changes in the input data. Although not specifically discussed in this report, analytical methods also have the potential to be optimized mathematically.

This chapter focuses on two types of reliability measures. Section 3.2 presents some existing network based methods for calculating measures involving the probability that an operational path exists between the sources and the demand nodes. These methods are then applied to some sample water distribution systems. Section 3.3 presents a network based method for calculating measures involving the probability of maintaining a specified flow at all demand points in a water distribution system. This method was developed by borrowing from similar methods for other types of networks.

3.2 REACHABILITY AND CONNECTIVITY

3.2.1 INTRODUCTION AND DEFINITIONS

In this chapter we will use the term "reachability" to denote the situation when *one specified* demand node in the network is connected to at least one source, and "connectivity" to denote the situation when *every* demand node in the network is connected to at least one source.

Reachability and connectivity are measures developed to quantify the reliability of electrical and communications networks. For water distribution systems (and other flow-carrying networks) connection to a source is only a necessary, not sufficient, condition to ensure that a given node is functional. For example, although a fully operational path may exist between a water source and a given demand node, if insufficient pressure exists in the system this demand node may not receive any water. Thus measures of connectivity and reachability are useful for performing an initial screening of the system. These measures can be used to quickly identify systems with serious problems due to insufficient redundancy to provide alternative routes when components fail. Similarly, demand nodes with reachability probabilities below those in the rest of the system may indicate serious supply shortages at these nodes.

For the calculation of these measures water distribution systems are often be modeled as networks of supply and demand nodes, connected by links. The nodes are modeled as being perfectly reliable. Each link i is said to have a probability p_i of functioning at any point in time and probability $q_i = 1 - p_i$ of failing.

Additionally, we also assume for all of the reliability and reachability methods discussed in this chapter that the links fail independently. Except for catastrophic occurrences (such as earthquakes) which are not accounted for by these methods, there is no reason to suspect significant interactions between pipe failures in one pipe and the future operation of others.

At any point in time the system can assume one of a large number of configurations, with some of the links functioning and other links failed. The probability of any one configuration occurring can be calculated as the product of the p_i's for the operative links times the product of the q_i's of the failed links. For connectivity calculations, each configuration corresponds to either a connected system, where *every* demand node is connected via functioning links to some source, or a disconnected system. Similarly for reachability, for a specified node each configuration corresponds to either a system in which the node is reachable via functioning links or a system in which the node is unreachable. Water distribution networks have the property of being "coherent", meaning that failure of an additional link can not cause a disconnected system to become connected, nor can it cause an unreachable node to become reachable.

Conceptually, calculating the overall probability of a given system being connected or the probability of a given node being reachable is a straightforward combinatoric problem. For any system, these probabilities can be calculated by testing each configuration individually, and adding up the probabilities of each configuration that is connected or for which a given node is reachable. However, there are an exponential number of configurations. Even for a 20 link system 2^{20}, or over one-million configurations, would have to be tested. Thus methods for calculating these probability measures must find efficient ways of searching and counting connected configurations.

It can be shown (Valiant [1979], Provan and Ball [1983a]) that calculating the connectivity and reachabilities for a general network is NP-hard. (NP-hard means that it is unlikely any polynomial time algorithms for these calculations exist.) However, classes of graphs can be found for which polynomial time algorithms for reachability and connectivity calculations can be developed. In particular, the next section will focus on existing polynomial-time methods for calculating reachability and connectivity on the class of series-parallel graphs. For more general networks, many algorithms exist which are exponential but are exponential to some base less than the number of links. Thus the key to calculating reachability and connectivity measures for more general networks involves finding some special structure that can be exploited in the reliability calculations. One such method which is likely to be useful for many general water distribution networks will be discussed in Section 3.2.3.

In order to discuss series-parallel graphs it is necessary to present a few definitions. This chapter will follow the notation and definitions relating to series-parallel graphs used by Satyanarayana and Wood [1982]. A series-parallel network is an undirected network reducable to a tree by performing only series and parallel reductions. A series reduction (Figure 3-1) can be performed by replacing two links (u – v and v – w), incident to the same node of degree 2 (node v) by one link (u – w). If the probability of operation of each original link is p_1 and p_2 respectively, the probability of operation of the new link will be p_1p_2. A parallel reduction (Figure 3-2) can be performed by replacing two links connecting the same two nodes by one link. If the probability of operation of each original link is p_1 and p_2 respectively, the probability of operation of the new link will be $1 - q_1q_2$. The overall probability that the network will remain connected, or that any node (except v) will remain reachable is not affected by these reductions.

A polynomial-time algorithm by Satyanarayana and Wood [1982] calculates the "K-terminal reliability" of an undirected series-parallel graph. For these

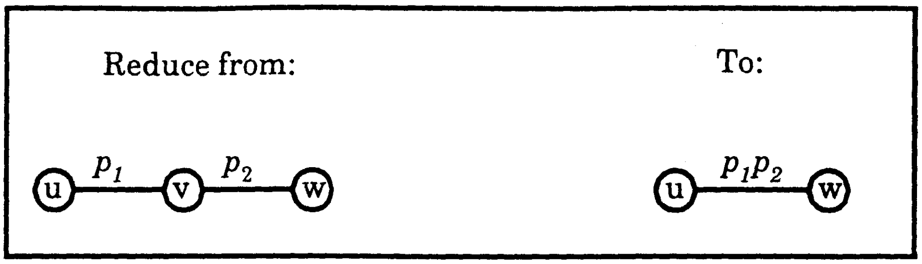


Figure 3-1 Series Reduction

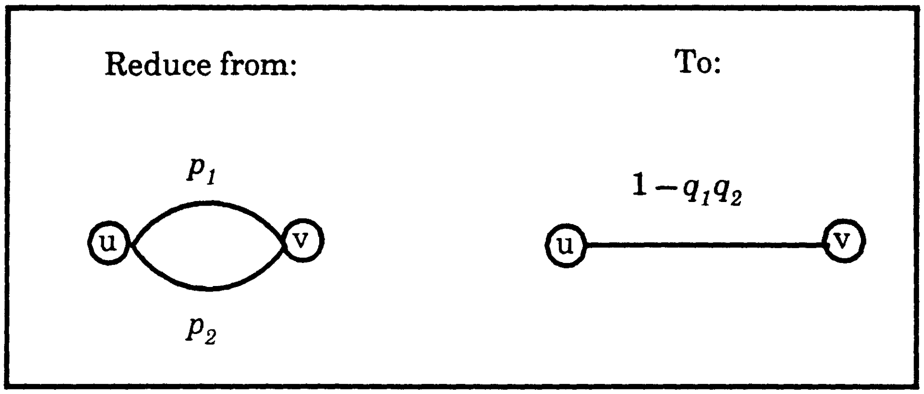


Figure 3-2 Parallel Reduction

calculations the analyst specifies a subset of nodes of interest in the graph; these special nodes are called K-vertices. The "K-terminal reliability" is the probability that all nodes in set K can communicate with one another (possibly through nodes not in set K). For set K specified as all nodes, the "K-terminal reliability" corresponds to connectivity as described in this section. For set K specified as two nodes, one a source and one a demand node, the "K-terminal reliability" corresponds to reachability for that node in a single-source system. Since any one node will be connected whenever the entire system is connected, it is obvious that the reachability for any node will always be greater than or equal to the connectivity for the graph as a whole.

Since K-vertices are in some sense "special", some restrictions on series and parallel reductions for components involving K-vertices exist. To maintain the correct reachability or connectivity for a graph with K-vertices, the series reduction given in Figure 3-1 is possible only when either: (1) node v *is not* a K-vertex or (2) *all three* nodes *are* K-vertices. If node v is not a K-vertex, the reduction in Figure 3-1 can be performed. When all three nodes are K-vertices the following adjustment is made:

(1) the probability of operation of the new link (u – w) becomes $p_1 p_2/(1 - q_1 q_2)$, and
(2) the system reliability correction factor is multiplied by $(1 - q_1 q_2)$.

After the graph has been reduced to a single edge, the reliability of the network is found by multiplying the probability of operation of the final edge by the reliability correction factor. Notice, if the series of all K-vertices reduction is the last reduction to be made, the reduction is equal to the non-K-vertex series reduction. Otherwise, this new reduction method accounts for the necessity that the "special" K-vertices be connected to the others, even though when the reduction is made some K-vertices nodes seem to "vanish".

Note, for a series-parallel graph with a single source s and a single terminal t identified, it may not be possible to reduce the graph via only series and parallel reductions to an edge $s - t$. Similarly, for a given set K of identified nodes, it may not be possible to reduce the series-parallel graph to a tree containing only those nodes. Satyanarayana and Wood [1982, p. 6] call series-parallel graphs which cannot be reduced to a tree of K-vertices by series and parallel reductions "complex series-parallel" graphs. A graph G_K which can be reduced to a tree of K-vertices by only series and parallel reductions is called "series-parallel" reducible. For example in Figure 3-3, if node a is chosen as the source node and node c as the terminal node ($K = \{a,c\}$), the graph G_K is series-parallel reducible.

In contrast, if b is chosen as the source node and d as the terminal node (K' = $\{b,d\}$), the graph G_K is series-parallel complex.

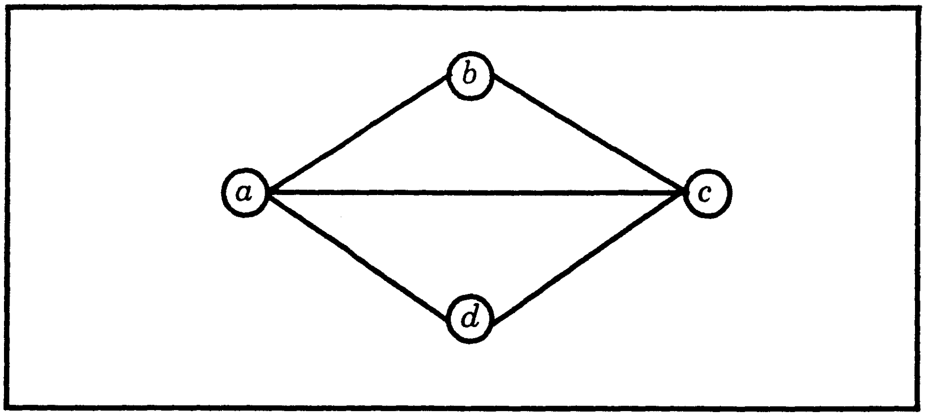


Figure 3-3 Series-Parallel Graph

All of the above definitions apply only to undirected graphs. For a directed graph D, an associated graph G called the underlying graph can be specified. The graph G will have the same nodes as D, but for each directed link in D, graph G will have an undirected link connecting the same nodes. Figure 3-4 presents a directed graph with the underlying graph of Figure 3-3. A directed graph is said to be basic series-parallel if its underlying graph is series-parallel.

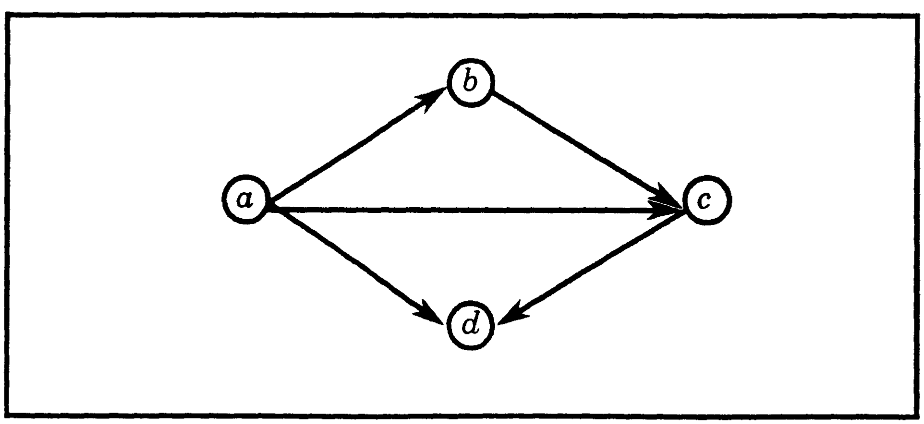


Figure 3-4 Directed Series-Parallel Graph

3.2.2 CALCULATION OF REACHABILITY AND CONNECTIVITY FOR SERIES-PARALLEL NETWORKS

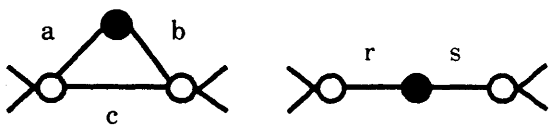
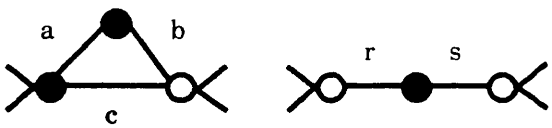
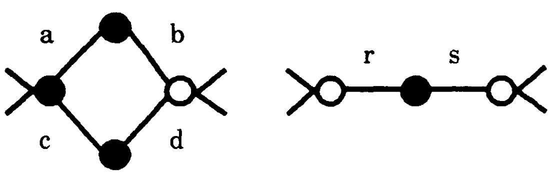
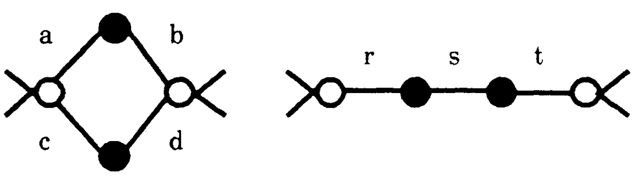
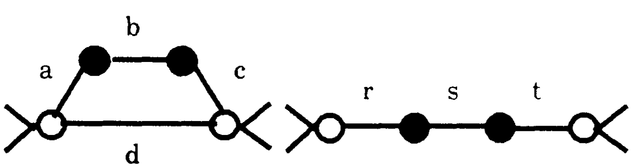
Methods for calculating connectivity for series-parallel networks can be usefully applied to the analysis of water distribution networks. Some, although certainly not all, water distribution networks are series-parallel. Series-parallel networks may be found in smaller systems which can be served by networks which are basically serially connected trees with a few connections added on. Also, a useful first step for some complex systems may be to abstract a series-parallel network from the general network, since many reliability calculations are considerably easier to do for series-parallel networks than for general networks. Reachability calculations for series-parallel networks involve the added constraint that the network involve only a single source. However, this constraint can be frequently avoided by connecting the sources to a new (fictitious) "super source" by only series and parallel connections.

The development of polynomial-time algorithms for reliability and connectivity of series-parallel networks has occurred fairly recently. Satyanarayana and Wood [1982] claim the development of the first such algorithm for undirected series-parallel complex networks. A similar algorithm for reachability on directed series-parallel complex networks was developed by Agrawal and Satyanarayana [1984]. Both of these methods are based on performing reductions which preserve the "K-terminal reliability" of the original network. The object of these reductions is to eventually reduce the graph to a single edge with reliability equal to that of the original system. This subsection will focus on the method for undirected networks since usually pipes and fittings can carry water in either direction, and thus can be represented by undirected links in a network.

This method proceeds by first making as many series and parallel reductions as possible. For a series-parallel graph, if simple series and parallel reductions do not reduce the system to a simple tree of K-vertices, the reduced network will always contain one of the polygons shown in Figure 3-5. The polygons can be reduced to a series connections of two or three links, as shown in the figure. Similarly to the previously described series reduction when all three nodes are K-vertices, (1) probabilities of operation for the "new" links are calculated and (2) factors are multiplied into the system reliability correction factor (Ω). In all cases, a polygon of three or more edges is reduced to two or three links connected serially.

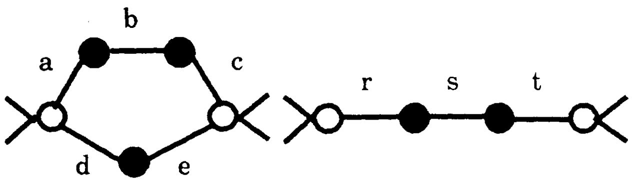
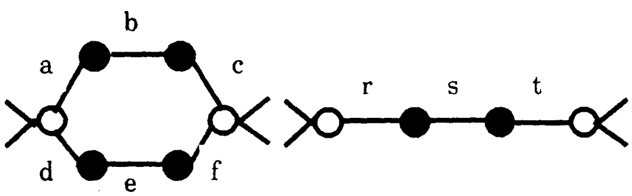
Once the polygons have been reduced, more series and parallel connections can be identified and reduced. By repeating the reduction process, the network will be reduced to a tree of only K-vertices, (usually a single edge). For a system reduced to a single edge, the system reliability equals the p_i for this edge multiplied by the correction factor. System reliability for trees is also easy to calculate, again the final reliability must be multiplied by the correction factor. Satyanarayana and Wood [1982] provide a formal algorithm for this process, and prove that the computational time involved is proportional to the number of links in the graphs; *e.g.* this is a linear-time algorithm.

As an example of the use of this algorithm, the connectivity and the reachability will be calculated for the series-parallel network, called Network A, with 10 nodes and 13 links (Figure 3-6). Network A represents a small system, with a reservoir at node 1 and a pumping station at link 100. Links 98 and 99 are pressure reducing valves, dividing the system into two pressure zones. Network A with 13 links has $2^{13} = 8192$ different combinations of operational and failed

Reduce polygon from: / To:	New Probabilities and Correction Factor:	Formulas:
	$p_r = \delta / (\alpha + \delta)$ $p_s = \delta / (\beta + \delta)$ $\Omega = (\alpha + \delta)(\beta + \delta) / \delta$	$\alpha = q_a p_b q_c$ $\beta = p_a q_b q_c$ $\delta = p_a p_b p_c \cdot (1 + q_a/p_a + q_b/p_b + q_c/p_c)$
	$p_r = \delta / (\alpha + \delta)$ $p_s = \delta / (\beta + \delta)$ $\Omega = (\alpha + \delta)(\beta + \delta) / \delta$	$\alpha = q_a p_b q_c$ $\beta = p_a q_b q_c$ $\delta = p_a p_b p_c \cdot (1 + q_a/p_a + q_b/p_b + q_c/p_c)$
	$p_r = \delta / (\alpha + \delta)$ $p_s = \delta / (\beta + \delta)$ $\Omega = (\alpha + \delta)(\beta + \delta) / \delta$	$\alpha = p_a q_b q_c p_d + q_a p_b p_c q_d + q_a p_b q_c p_d$ $\beta = p_a q_b p_c q_d$ $\delta = p_a p_b p_c p_d \cdot (1 + q_a/p_a + q_b/p_b + q_c/p_c + q_d/p_d)$
	$p_r = \gamma / (\alpha + \gamma)$ $p_s = \gamma / (\beta + \gamma)$ $p_t = \gamma / (\delta + \gamma)$ $\Omega = (\alpha + \gamma)(\beta + \gamma) \cdot (\gamma + \delta) / \gamma^2$	$\alpha = q_a p_b q_c p_d$ $\beta = p_a q_b q_c p_d + q_a p_b p_c q_d$ $\delta = p_a q_b p_c q_d$ $\gamma = p_a p_b p_c p_d \cdot (1 + q_a/p_a + q_b/p_b + q_c/p_c + q_d/p_d)$
	$p_r = \gamma / (\alpha + \gamma)$ $p_s = \gamma / (\beta + \gamma)$ $p_t = \gamma / (\delta + \gamma)$ $\Omega = (\alpha + \gamma)(\beta + \gamma) \cdot (\gamma + \delta) / \gamma^2$	$\alpha = q_a p_b p_c q_d$ $\beta = p_a q_b p_c q_d$ $\delta = p_a p_b q_c q_d$ $\gamma = p_a p_b p_c p_d \cdot (1 + q_a/p_a + q_b/p_b + q_c/p_c + q_d/p_d)$

Note: Solid nodes represent K-vertices

Figure 3-5 Polygon-to-Chain Reductions (Adapted from Satyanarayana and Wood [1982, p. 8])

Reduce polygon from:	To:	New Probabilities and Correction Factor:	Formulas:
		$p_r = \gamma / (\alpha + \gamma)$ $p_s = \gamma / (\beta + \gamma)$ $p_t = \gamma / (\delta + \gamma)$ $\Omega = (\alpha + \gamma)(\beta + \gamma) \cdot (\gamma + \delta) / \gamma^2$	$\alpha = q_a p_b p_c q_d p_e$ $\beta = p_a q_b p_c \cdot (p_d q_e + q_d p_e) + p_b (q_a p_c p_d q_e + p_a q_c q_d p_e)$ $\delta = p_a p_b q_c p_d q_e$ $\gamma = p_a p_b p_c p_d p_e \cdot (1 + q_a/p_a + q_b/p_b + q_c/p_c + q_d/p_d + q_e/p_e)$
		$p_r = \gamma / (\alpha + \gamma)$ $p_s = \gamma / (\beta + \gamma)$ $p_t = \gamma / (\delta + \gamma)$ $\Omega = (\alpha + \gamma)(\beta + \gamma) \cdot (\gamma + \delta) / \gamma^2$	$\alpha = q_a p_b p_c q_d p_e p_f$ $\beta = p_a q_b p_c \cdot (q_d p_e p_f + p_d q_e p_f + p_d p_e q_f) + p_a p_b q_c p_f (p_d q_e + q_d p_e) + q_a p_b q_c p_d (q_e p_f + p_e q_f)$ $\delta = p_a p_b q_c p_d p_e q_f$ $\gamma = p_a p_b p_c p_d p_e p_f \cdot (1 + q_a/p_a + q_b/p_b + q_c/p_c + q_d/p_d + q_e/p_e + q_f/p_f)$

Note: Solid nodes represent K-vertices

Figure 3-5 Polygon-to-Chain Reductions (continued)

nodes, yet the connectivity can be calculated with only a few pages of calculations.

Figure 3-7 presents a summary of the connectivity calculations for Network A. Connectivity can be calculated by performing only series and parallel reductions. However, since connectivity corresponds to the case where all nodes are

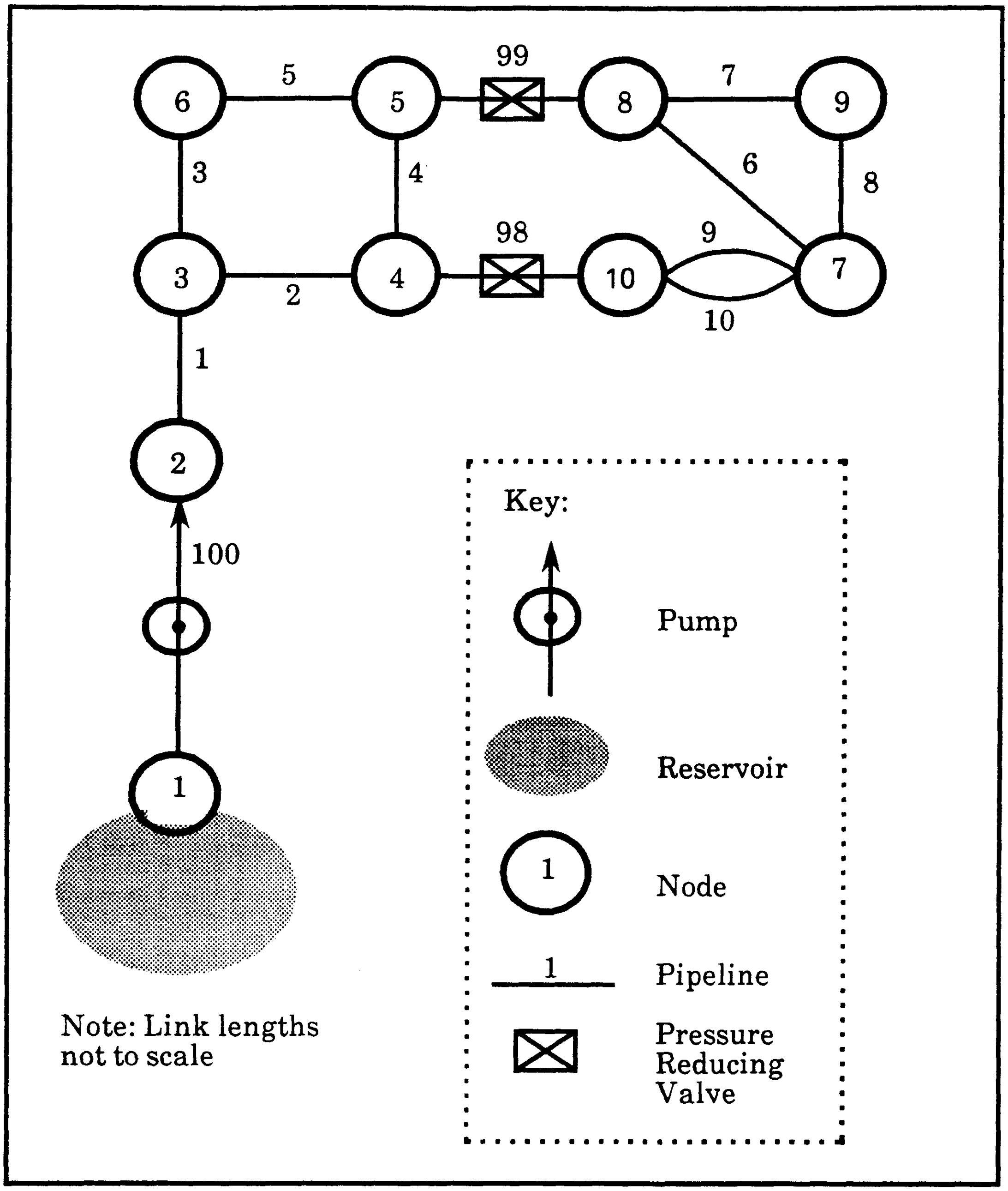


Figure 3-6 Network A (with link numbers)

K-vertices, for all series reductions the new edge probabilities must be adjusted as described in the last section, and a correction factor calculated. Figure 3-8 presents a summary of the reachability calculations for node 7. Reachability calculations for this node involve using the first of the polygon reductions given in Figure 3-5 (in Figure 3-8 this corresponds to the fourth reduction). All other reductions for node 7 calculations are series or parallel reductions. Table 3-1 presents the reachability formulas for each node.

To give a more intuitive understanding of this system, Table 3-2 gives numerical values for the connectivity of the system and the reachability of each node when:

* all links have $p_i = 0.95$
* all links have $p_i = 0.97$
* all links have $p_i = 0.99$
* all links have $p_i = 0.999$
* link 1-2 has $p_i = 0.9543$ (pump) and the rest have $p_i = 1 - (1.557 \times 10^{-6}) \times$ pipe length (ft).

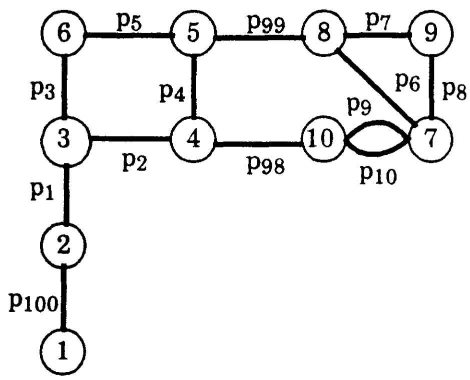
The rationale for the probability values in the last case is as follows. Link 100 is a pump, which is assumed to fail on average 8 times per year for 50 hours each time. (These values are in the range of those given in Damelin, *et. al.* [1972].) Thus the pump is assumed to be operational with probability $1 - (8 \cdot 50) / (24 \cdot 365)$. For the pipes, a breakage value of 1 break/ mile/ year (from O'Day [1982, p. 591] , see table 4-5) was chosen. To investigate the "worst-case" behavior of this system, this pipe breakage value is on the order of the highest value from data from 15 U.S. cities. We further assumed each break lasted 3 days. Thus each pipe is assumed to be operational with probability:

$$1 - (3 \cdot 24) / (24 \cdot 365) \cdot (\text{pipe length (ft)}/5280).$$

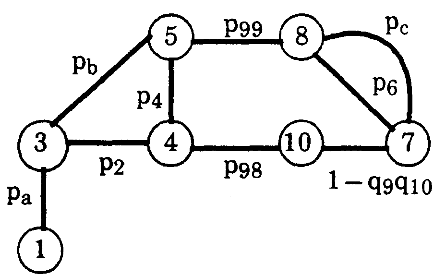
Table 3-3 presents the length and corresponding link probability of each link for this case. Note, although these connectivity and reachability values are reported only to four decimal places, as many digits as possible were carried in the computer for the calculation of these values.

Note: Every node is a K-vertex

Reliability Correction Factors

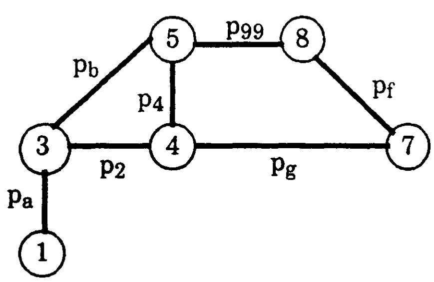


1.0



$(1-q_1q_{100})\,(1-q_3q_5)\,(1-q_7q_8)$

$$pa = p_1p_{100} / (1-q_1q_{100})$$
$$pb = p_3p_5 / (1-q_3q_5)$$
$$pc = p_7p_8 / (1-q_7q_8)$$

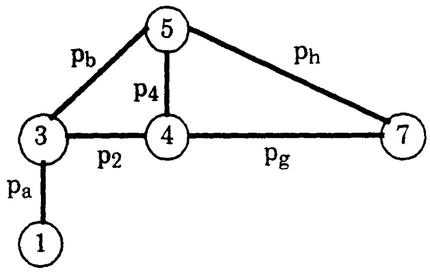


$(1-q_9q_{10}q_{98})$

$$pf = 1-q_6q_c$$
$$pg = p_{98}(1-q_9q_{10}) / (1-q_9q_{10}q_{98})$$

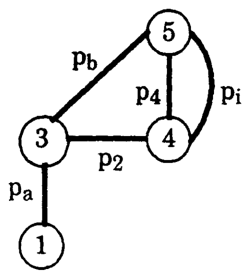
Figure 3-7 Connectivity Calculations for Network A

Reliability Correction Factors



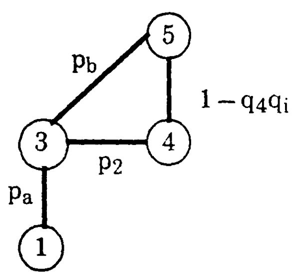
$p_h = p_{99}p_f / (1 - q_{99}q_f)$

$(1 - q_{99}q_f)$

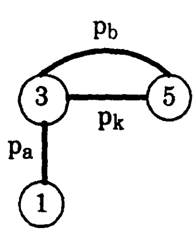


$p_i = p_g p_h / (1 - q_g q_h)$

$(1 - q_g q_h)$



1.0

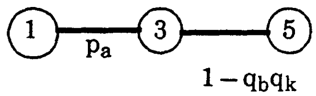


$p_k = p_2(1 - q_4 q_i) / (1 - q_2 q_4 q_i)$

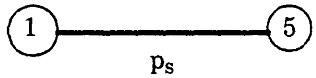
$(1 - q_2 q_4 q_i)$

Figure 3-7 Connectivity Calculations for Network A (continued)

Reliability Correction Factors



1.0



$(1-q_aq_bq_k)$

$p_s = p_a(1-q_bq_k)/(1-q_aq_bq_k)$

Connectivity $= p_s(1-q_aq_bq_k)(1-q_2q_4q_i)(1-q_gq_h)(1-q_{99}q_f)(1-q_9q_{10}q_{98})(1-q_1q_{100})(1-q_3q_5)(1-q_7q_8)$

Figure 3-7 Connectivity Calculations for Network A (continued)

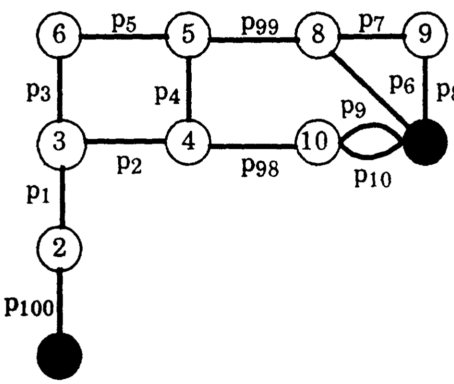
It is obvious that the reachabilities are very similar at each node. However, in the less reliable cases it can be seen that the nodes in the lower pressure zone (nodes 7,8,9, and 10) have slightly lower reachabilities, with node 9 the lowest of all. Since the reachabilities of these nodes are so alike, the connectivity value is a good summary measure for this reliability question. From the last case, with the varied link probabilities, it is obvious that the connectivity of this system is very close to the probability of operation of the first link – the pump. Indeed, as shown in Figure 3-9, for cases with equal link probabilities, the connectivity of Network A varies linearly with the p_i value.

3.2.3 CALCULATION OF REACHABILITY AND CONNECTIVITY FOR GENERAL NETWORKS

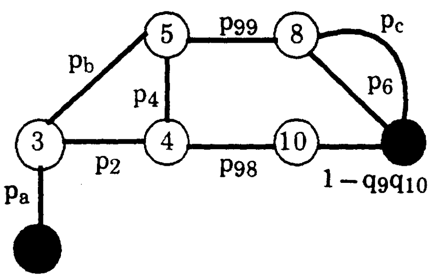
For non-series-parallel water distribution networks more complex algorithms must be used for the calculation of reachability and connectivity. As discussed in Chapter 2 there are several algorithms suitable for these calculations, each of

Note: Nodes 1 and 7 are K-vertices

<u>Reliability Correction Factors</u>

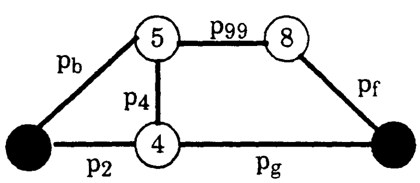


1.0



1.0

$pa = p_1 p_{100}$
$pb = p_3 p_5$
$pc = p_7 p_8$

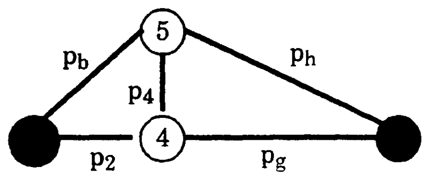


p_a

$pf = 1 - q_6 q_c$
$pg = p_{98}(1 - q_9 q_{10})$

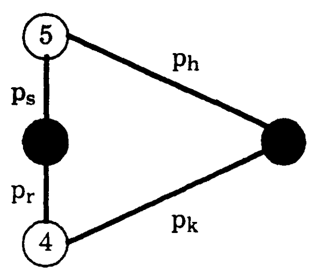
Figure 3-8 Reachability Calculations for Network A – Node 7

Reliability Correction Factors



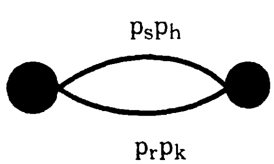
1.0

$p_h = p_{99} p_f$



$(\alpha + \delta)(\beta + \delta) / \delta$

$\alpha = q_2 p_b q_4$
$\beta = p_2 q_b q_4$
$\delta = p_2 p_b p_4 (1 + q_2/p_2 + q_b/p_b + q_4/p_4)$
$p_r = \delta / (\alpha + \delta)$
$ps = \delta / (\beta + \delta)$



1.0

1.0

$1 - (1 - p_s p_h)(1 - p_r p_k)$

Reachability of node 7 = $[1 - (1 - p_s p_h)(1 - p_r p_k)] [(\alpha + \delta)(\beta + \delta) / \delta] [p_a]$

Figure 3-8 Reachability Calculations for Network A – Node 7 (continued)

Node	Reachability
2	p_{100}
3	$p_{100}p_1$
4	$p_{100}p_1(1-q_2(1-p_3p_5(1-q_4q_a)))$
5	$p_{100}p_1(1-(1-p_3p_5)(1-p_2(1-q_4q_a))$
6	$p_{100}p_1(1-q_3(1-p_2p_5(1-q_4q_a)))$
7	$R(1-(1-p_sp_{99}p_t)(1-p_rp_{98}p_u))$
8	$R(1-(1-p_sp_{99})(1-p_rp_{98}p_up_t))$
9	$R \cdot R'(1-(1-p_sp_{99}p_v)(1-p_rp_{98}p_up_w))$
10	$R(1-(1-p_rp_{98})(1-p_sp_{99}p_tp_u))$

where:

$p_a = p_{99}p_{98}(1-q_6(1-p_7p_8))(1-q_9q_{10})$

$q_a = 1 - p_a$

$p_n = p_3p_5$

$\alpha = q_2p_nq_4$

$\beta = p_2q_nq_4$

$\delta = p_2p_np_4(1+ q_2/p_2 + q_n/p_n + q_4/p_4)$

$R = p_{100}p_1(\alpha + \delta)(\beta + \delta)/\delta$

$p_r = \delta / (\alpha + \delta)$

$p_s = \delta / (\beta + \delta)$

$\alpha' = q_7p_8q_6$

$\beta' = p_7q_8q_6$

$\delta' = p_7p_8p_6(1+ q_7/p_7 + q_8/p_8 + q_6/p_6)$

$R' = (\alpha' + \delta')(\beta' + \delta')/\delta'$

$p_v = \delta' / (\alpha' + \delta')$

$p_w = \delta' / (\beta' + \delta')$

$p_u = 1-q_9q_{10}$

Table 3-1 Reachability Formulas for Network A

NETWORK A

Case I: $p_i = 0.95$ for all i

Case II: $p_i = 0.97$ for all i

Case III: $p_i = 0.99$ for all i

Case IV: $p_i = 0.999$ for all i

Case V: p_i from Table 3-3

Connectivity:

I	II	III	IV	V
0.8902	0.9364	0.9796	0.9980	0.9540

Reachability:

Node	I	II	III	IV	V
2	0.9500	0.9700	0.9900	0.9990	0.9543
3	0.9025	0.9409	0.9801	0.9980	0.9540
4	0.8979	0.9392	0.9799	0.9980	0.9540
5	0.8977	0.9391	0.9799	0.9980	0.9540
6	0.8979	0.9392	0.9799	0.9980	0.9540
7	0.8952	0.9382	0.9798	0.9980	0.9540
8	0.8952	0.9382	0.9798	0.9980	0.9540
9	0.8930	0.9374	0.9797	0.9980	0.9540
10	0.8952	0.9382	0.9798	0.9980	0.9540

Table 3-2 Connectivity and Reachability Results for Network A

Link	Length (ft)	p_i
1	200	0.9997
2	1500	0.9977
3	1800	0.9972
4	2000	0.9969
5	1900	0.9970
6	1000	0.9984
7	2500	0.9961
8	3500	0.9946
9	1500	0.9977
10	1500	0.9977
98	500	0.9992
99	500	0.9992
100	pump	0.9543

Table 3-3 Network A Link Lengths and Probabilities

which was designed for networks with different special structures. Thus it is impossible to choose one algorithm which is most useful for general water distribution networks. However, most water distribution networks are at least (nearly) planar, *i.e.* they can be drawn on a plane with (little or) no crossing of lines. Planarity limits the number of possible interconnections, and thus water distribution networks are ensured to be relatively sparse.

Rosenthal [1977] presents an algorithm which can be used to calculate a number of reliability measures on complex networks. Although in the worst case the total computational effort required by this algorithm grows more than polynomially with the problem size, for "treelike" and other sparse networks the total computational effort grows only linearly. Thus Rosenthal's algorithm should be

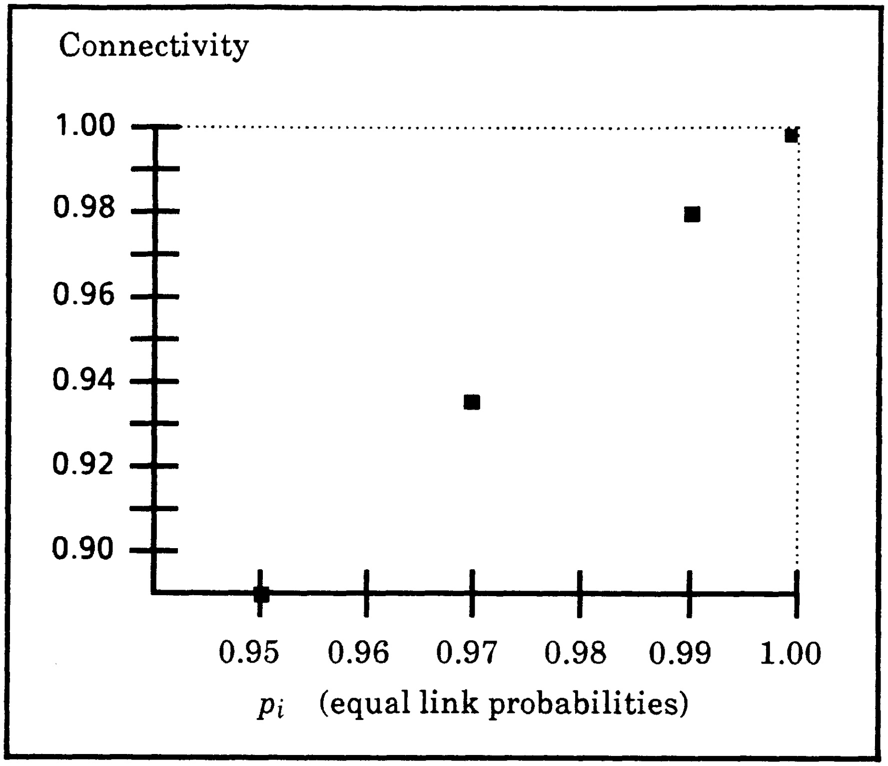


Figure 3-9 Graph of Connectivity versus Link Probability for Network A

useful for reliability calculations for many water distribution systems. Rosenthal describes in detail how this algorithm can be used to calculate a number of reliability measures, including reachability, connectivity, and the K-terminal reliability described in the previous subsection. The algorithm as used for connectivity calculations is outlined below, followed by the calculation of the connectivity of a sample non-series-parallel network.

Rosenthal's algorithm is used with a model of the same type as in the previous subsection, namely a network with perfectly reliable nodes, statistically independent link failures, and with the probability of operation of each link given by p_i. In this method, the network is decomposed into a number of subnetworks. Subnetworks are classified by the number of boundary nodes they contain. A boundary node is any node incident to an arc connecting the subnetwork to the larger network. In Figure 3-10, a subnetwork is circled. Nodes 70 and 100 are boundary nodes for this subnetwork.

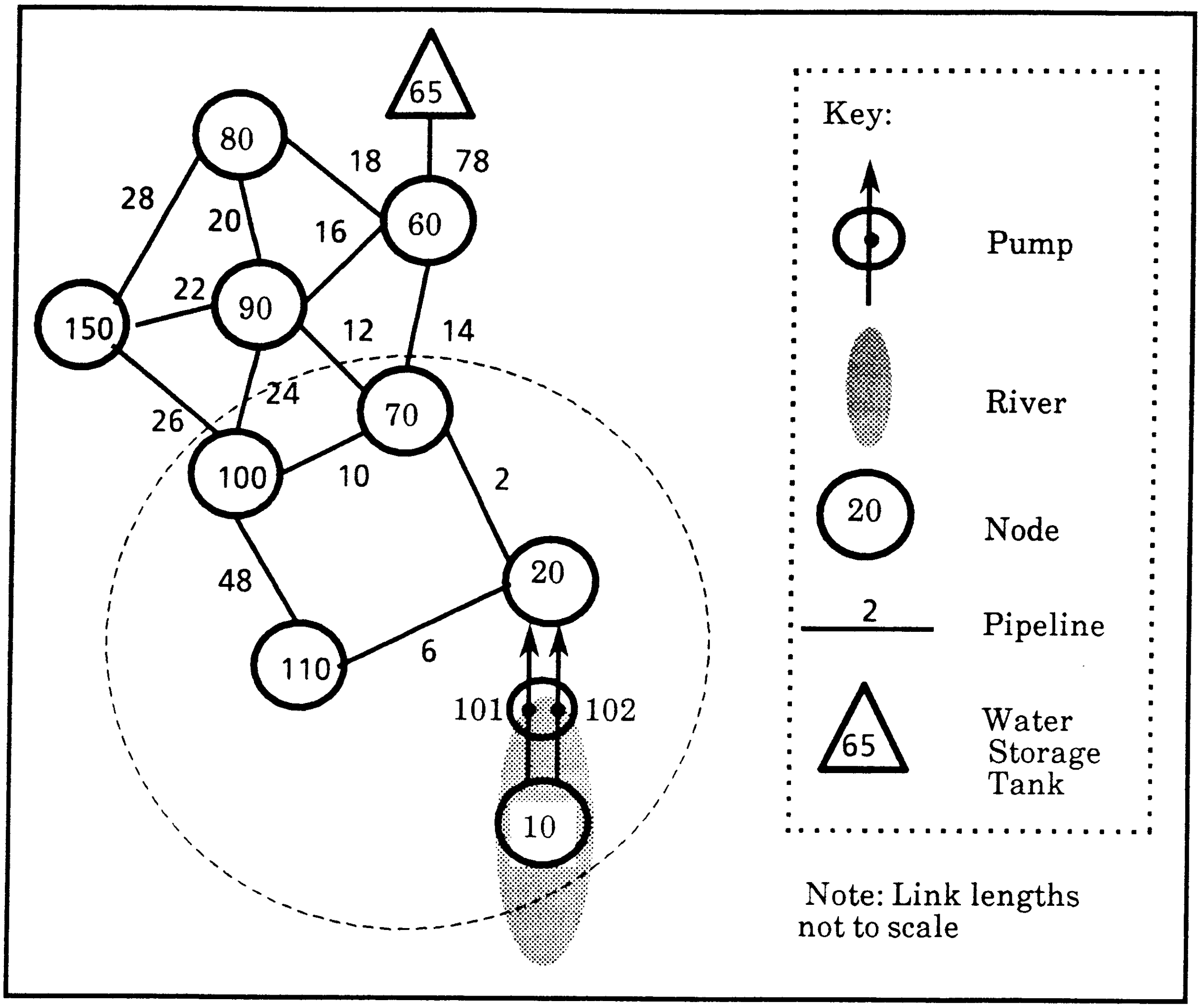


Figure 3-10 Network B

For a subnetwork with two boundary nodes (u and v), called a 2-subnet, the possible configurations of this subnetwork can be divided into three separate classes (Murchland [1957]):

"failed" (denoted sf) = states in which nodes are isolated from the boundary nodes
[u,v] = states not in sf for which u communicates with v via the 2-subnet
[u][v] = states not in sf for which u does not communicate with v via the 2-subnet.

Classes for subnetworks with more than two boundary nodes can also be identified by determining which subsets of the boundary nodes can communicate via the subnetwork and which cannot. Each possible subset grouping is a separate class. For a 3-subnet with boundary nodes (u,v, and w) the following six classes exist: sf, [u,v,w], [u][v,w], [v][u,w], [w][u,v], [u][v][w]. For example, [u][v,w] is the class of configurations for which u is not connected to v and w in the subnetwork. However, u may communicate with v and w through the larger network, if proper connections in the larger network are operative.

The algorithm is initialized by dividing the network so that each link is a separate subnetwork. The algorithm proceeds by combining in each step, two subnetworks with at least one boundary node in common. For the larger subnetwork the probability of it being in each of the possible classes is then calculated. The algorithm ends when all the subnetwork have been combined into one network. Then the probability of being in class [u,v,w, . . . , z] is the connectivity for the entire network.

From Rosenthal, it is easily proven that when combining two subnetworks (S_1 and S_2) into a larger network (S_3) the composition of any two classes (C_1 from S_1 and C_2 from S_2), yields a unique class of the larger subnetwork. Thus the probabilities associated with the new classes are calculated with the following formula:

$$Pr(C_3) = \Sigma\, Pr(C_1 \text{ and } C_2) = \Sigma\, Pr(C_1)Pr(C_2) \tag{3-1}$$

where the sum is over all classes (C_1 and C_2) which combine to produce C_3.

Rosenthal does not specify the order in which the subnetworks should be combined. In fact he indicates he "tried unsuccessfully to find a good algorithm to determine an optimal sequence of subnetwork merges to minimize ... [the] total computational work". In one sets of computational experiments on variations of a computer communications network linking some U.S. universities (1974 ARPANET), Rosenthal used the heuristic rule: first the links in series or parallel configurations were combined, followed by successively merging 2-subnets into one growing subnetwork. He was able to solve a 46 node/ 63 arc problem with this algorithm, and needed to employ only subnetworks of 6 boundary nodes or less.

One iteration of this algorithm will be illustrated in the calculation of the connectivity of Network B, an 10 node, 16 link network (shown in Figure 3-10). Network B is a portion of a test problem "Anytown" proposed by Thomas Walski for analysis at the June 1985 ASCE Water Resources Conference in Buffalo, NY. This small network, a subset of the Anytown network, represents the central part of the city. A subset of the network was chosen to allow calculation by hand. This network is not series-parallel, for example the pentagon of nodes and links surrounding node 90 cannot be reduced at all with series or parallel reductions.

In Network B, node 10 represents a water treatment plant, node 65 represents a water storage tank, and the two links from node 10 to 20 represent pumps. Figure 3-11 presents a summary of a few steps of the connectivity calculations for this network. Again, for a better understanding of this reliability expression, Table 3-4 presents the numerical value of this expression when:

* all links have $p_i = 0.95$
* all links have $p_i = 0.97$
* all links have $p_i = 0.99$
* all links have $p_i = 0.999$
* link 1-2 has $p_i = 0.9543$ (pump) and the rest have $p_i = 1 - (1.557 \times 10^{-6}) \times$ pipe length (ft).

The last set of p_i's again comes from a figure of 1 break/mile/year. Table 3-5 presents the length and corresponding link probability of each link in Network B for this case.

Although this network is more complex than network A the connectivities are quite similar. Again, as shown in Figure 3-12, for cases with equal p_i's the connectivity of Network B varies linearly with the p_i value. The connectivity calculated from the last probability set is again very close to the probability of operation of the pump, despite the possible modifying influence of the water storage tank.

Because the nodes are modeled as perfectly reliable, the connectivity expression calculated above is actually an upper bound for the reliability of this system. Node 65 represents a water tank, thus the assumption of perfect node reliability implies the tank contains an infinite volume of water. Even if unreliable nodes are allowed in the analysis, the assumption of independence of failures does not allow modeling the length of time until the tank runs out of water as dependent upon on the current configuration of the system. To model this more complex (but more realistic) situation, fault-tree methods or simulation methods are needed. Water tanks with finite volumes are modeled in the simulations in Chapter 4.

Assume the subnetwork identified in Figure 3-10 (nodes 10,20,70,100, and 110) has been solved. Nodes 70 and 100 are the boundary nodes. Let the following represent the probabilities of each class:

Class	Probability
sf	a_1
[70,100]	a_2
[70][100]	a_3

Assume another subnetwork (nodes 60,65, and 70) has also been solved. Nodes 60 and 70 are the boundary nodes. Let the following represent the probabilities of each class:

Class	Probability
sf	b_1
[60,70]	b_2
[60][70]	b_3

These subnetworks can be merged into a 3-subnet, consisting of nodes 10, 20, 60, 65, 70, 100, 110, 120. Nodes 60,70, and 100 are the boundary nodes. The probabilities of each class are:

Class	Probability
sf	$a_1 + a_2b_1 + a_3b_1$
[60,70,100]	a_2b_2
[60][70,100]	a_2b_3
[70][60,100]	0
[100][60,70]	a_3b_2
[60][70][100]	a_3b_3

Figure 3-11 One Iteration of Rosenthal's Algorithm Used to Calculate the Connectivity of Network B

NETWORK B

Case I: $p_i = 0.95$ for all i

Case II: $p_i = 0.97$ for all i

Case III: $p_i = 0.99$ for all i

Case IV: $p_i = 0.999$ for all i

Case V: p_i from Table 3-5

Connectivity:

I	II	III	IV	V
0.8956	0.9383	0.9798	0.9980	0.9534

Table 3-4 Connectivity and Reachability Results for Network B

3.2.4 DISCUSSION

As discussed previously, the major use of connectivity and reachability values is to provide a fast, easy to perform, check for reliability problems due to inadequate network interconnections or extremely unreliable links. For some systems, namely those that can represented by series-parallel networks, the algorithm of Satyanarayana and Wood may be used to calculate reachability and connectivity simply and extremely quickly. For more complicated networks, algorithms such as Rosenthal's exist for the calculations of these measures.

One possible problem with these methods is in determining the correct values for the probability of operation for each link. These values may be extrapolated from

Link	Length (ft)	p_i
101,102	pump	0.9543
78	100	0.9998
2	12000	0.9813
6	12000	0.9813
10	6000	0.9907
12	6000	0.9907
14	6000	0.9907
16	6000	0.9907
18	6000	0.9907
20	6000	0.9907
22	6000	0.9907
24	6000	0.9907
26	6000	0.9907
28	6000	0.9907
48	6000	0.9907

Table 3-5 Network B Link Lengths and Probabilities

historical records of the system. For example records may show the average pipe in a system is operational all but one day per year, giving a p_i of 364/365 = 0.9973. Similarly, as was done in the examples, these probabilities may be estimated to be related to the length of a link with longer links usually being less reliable.

No matter how these probabilities are estimated however, these methods always reflect the assumption that the probabilities are "instantaneous", *i.e.* they represent the probability of the link being operational *at any point in time.* This assumption implies that instants of time are independent. In these methods a

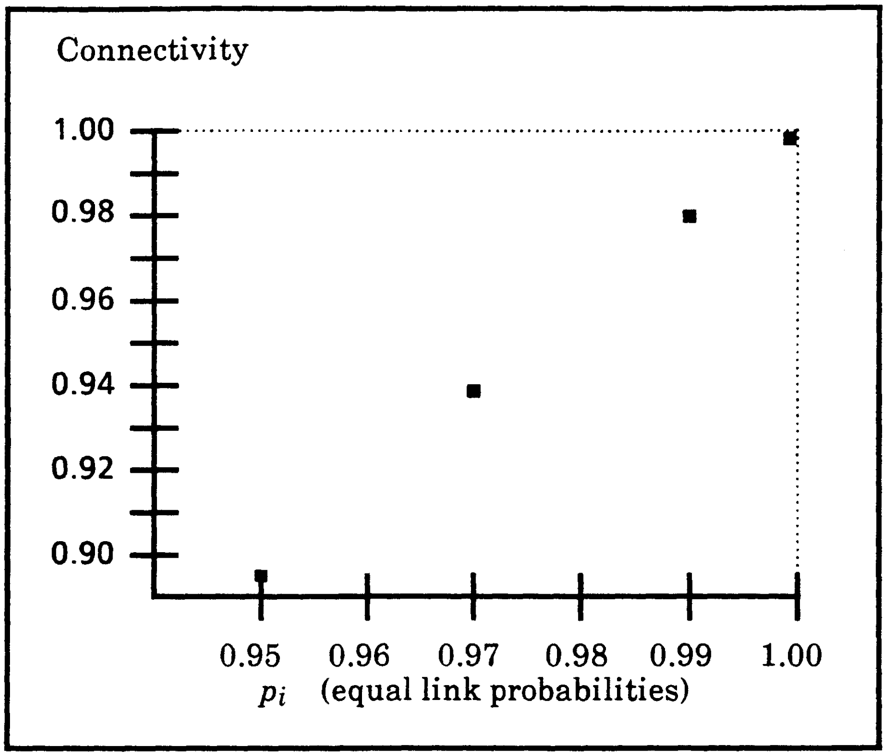


Figure 3-12 Graph of Connectivity versus Link Probability for Network B

link that was inoperative for one ten-hour period per one year would have reliability equal to a link that was down for one hour ten times during the year. Of course, once a pipe has failed it will stay failed for at least a few hours. Thus the probability that a pipe is failed at some moment in time is not completely independent of it's operational state the moment before. This model also cannot exactly represent a situation where, for example, a pipe failure lasts no more than one day, no matter how long the pipe has previously been operational. For use as a fast first step in a reliability assessment however, this "point in time" assumption is unlikely to be of much concern.

3.3 PROBABILISTIC SUPPLY MEASURES

3.3.1 INTRODUCTION AND DEFINITIONS

As discussed in the previous section, reachability is not sufficient to establish that the supply at a demand node is being met. For a reachable node, it is important the node receive sufficient supply at adequate pressure. Receiving adequate supply is a function not just of the arrangements of links in the network, but also of the amount of flow that can be carried along these links.

Over time, as links randomly fail and are fixed, the amount of supply or the pressure of that supply will vary at any given node. Thus it would be useful to know for each node the probability distribution of these quantities. Unfortunately, methods for calculating these distributions, or even the mean and variances of these distributions can only handle very simple systems with no more than a few nodes. There are, however, network based methods that can calculate measures such as "the probability that x units of supply per minute can be delivered at a given node", or "the probability that each node i can be supplied with x_i units of supply per minute at a specified pressure".

Conceptually, calculating the above probabilistic supply measures is again a simple combinatoric problem, almost exactly like the problem of calculating network connectivity and reachabilities. Again, the water distribution system can be modeled as a network with perfectly reliable nodes, statistically independent link failures, and each link i having a probability of failing (p_i) at any point in time. Then the probability that all nodes have sufficient supply can theoretically be calculated by adding up the probability of each individual configuration for which the system can provide the required supply. As with connectivity methods, the key to finding efficient algorithms is to find efficient ways to search for operational configurations.

Testing for sufficient supply in a network is however much more difficult than testing for connectivity of that network. Hydraulic networks behave according to a simultaneous set of non-linear algebraic equations. Computer programs for solving for the pressures and supplies in hydraulic networks do exist, but they employ iterative techniques. It is also impossible to use the solution of the full system to quickly deduce the solution to a reduced system, since flows can be rerouted in complex ways. Even with an efficient algorithm for searching among system configurations, a large number of configurations must be tested requiring considerable amounts of computer time. Thus, in order to calculate these probabilistic supply measures, some easily solved approximation of the hydraulic characteristics of the full network must be found.

In this section, the use of a capacitated network as a model of the full hydraulic network will be explored. A capacitated network model simply adds, to the previously described network model, maximum capacities on each arc. Now, the problem of determining if a given configuration can meet the specified demand without exceeding the capacities of each link has been reduced to a classical network flow problem, usually called a transshipment problem. Many well-known polynomial-time algorithms exist for finding a feasible solution for a transshipment problem for *one* network. However, it is a much more difficult problem to find the *probability* that a network with unreliable links will operate, which involves testing *many* individual network configurations. It can be proved (Valiant, [1979]) that even for series-parallel networks the problem of finding the probability that each node will receive sufficient supply is NP-hard.

In the following subsections, algorithms for calculating probabilistic supply measures on capacitated networks will be discussed, methods for assigning the arc capacities will be developed, and these measures calculated for two sample networks. Based on these examples, the validity of using a capacitated network supply model to represent a water distribution system will be discussed.

3.3.2 CALCULATION OF PROBABILISTIC SUPPLY MEASURES ON CAPACITATED NETWORKS

The objective of this section is to calculate the probability that every node in the network receives sufficient water supply. This probabilistic supply measure is similar to connectivity, except for this measure configurations in which all nodes are connected may be still be infeasible due to lack of sufficient flow-carrying capacity.

The development of methods for calculating the reliability of capacitated networks has not received as much attention in the literature as has development of methods for calculating connectivity. One of the conceptually easiest methods was developed by Lee [1980]. Lee's algorithm, based on lexicographic ordering, provides the search strategy among the possible configurations. At each iteration this algorithm determines a set of operational configurations and accounts for the probabilities of this feasible set. Lexicographic ordering in an extension of the principle that the last arc on the list is deleted first.

In the worst case and with a poor choice of initial configuration, this algorithm may search every one of the 2^L configurations (where L is the number of links in the network). However, for sparsely connected networks, and networks that are not highly reliable, this algorithm is fairly efficient. Thus for a first investigation of the use of capacitated networks Lee's algorithm was employed to compute the probability that the system can meet the specified demand at each node without exceeding the capacity of each arc.

The steps involved in using this algorithm are:

1) finding an initial feasible solution to the capacitated network, namely a solution that supplies the specified demand at each node without exceeding the capacity of any arc,

2) successively removing operational links and seeing if a feasible solution to the reduced network can be found,

3) accumulating the probabilities of feasible configurations, and

4) repeating steps 2) and 3) for all links which participate in any feasible solution until all feasible configurations have been accounted for.

For ease in solving the transshipment problem for the initial and the reduced networks, the network is augmented as follows:

* a node representing a "super source" is added with supply equal to the total network demand

* a link is added from the "super source" to each supply node with capacity equal to the maximum supply at that node

* each undirected link i-j in the network is replaced by two directed links ($i \rightarrow j$ and $j \rightarrow i$) each with capacity equal to the original link.

The classical spanning-tree based transshipment algorithm (see for example Bradley, Hax and Magnanti, [1977, p. 326-335]) was used to solve the individual network configurations. The solution to the reduced network can be easily found as the minimum cost solution to the original network, when the costs on the arcs are as follows:

* the cost of each failed arc is 1 unit

* the cost of each operational arc is 0 units.

(Of course, a failed link i-j implies that both directed arcs $i \rightarrow j$ and $j \rightarrow i$ have failed.)

If the minimum cost solution to the above problem has zero cost, a feasible solution using only operational arcs is possible. If the minimum cost solution is greater than zero, no solution is possible without using failed arcs indicating that the reduced system can not meet the specified demand. The feasible flows carried on each link, when every link is operating, are supplied as an initial feasible solution to the transshipment problem for the reduced networks.

The computer program used for the analysis of the following sample systems employed Lee's algorithm to search among configurations, linked to the classical transshipment algorithm for testing the individual configurations. In this program the initial feasible solution to the full network is required as a starting point. It was not necessary to implement a search for the initial feasible solution, because for networks of the size that can be handled by Lee's algorithm it is a simple matter to find the initial feasible solution by hand.

As an example of the use of this algorithm, the Networks A and B (Figures 3-6 and 3-10) from the previous section were analyzed. The following paragraphs will discuss details of the probabilistic supply analysis. Suitability of the model will be discussed in the next subsection.

It is not clear how to determine the capacity of each link. Most pipes are rated as to the maximum pressure they can withstand, so one limit to the amount of flow they can carry may be obtained by setting the pressure at one end of that pipe to the maximum allowable pressure, setting pressure at the other end to 0, and solving for the flow carried by the pipe. Certainly this value will be an upper bound on the capacity of the pipe, however it is unlikely that there will ever be sufficient pressure in the system to obtain this limit. In most systems, the flow through each link will be determined by the number and location of the pumps, elevated water tanks, and other pressure modifying devices, rather than by the pipe pressure limits. It is difficult, however, to determine how the pumping capacity of each pump and the elevation of each tank will affect the flow through some pipe not directly connected to that pump or tank.

As an attempt to avoid very complicated procedures, but to still reflect practical limits on the flow in each pipe, a "rule of thumb" estimation was tried, namely the maximum flow in each pipe will be the flow that would occur if the pipe was installed at a gradient of 0.01. Thus the maximum flow in each link was given by the Hazen-Williams equation (see for example Walski [1984a, p. 35]):

$$V = 0.55 \times C \times D^{0.63} \times S^{0.54} \tag{3-2}$$

From (3-2) we get:

$$Q = 0.2795 \times V \times (\pi \times D^2)/4 = 4.057 \times 10^{-4} \times C \times D^{2.63} \times S^{0.54} \tag{3-3}$$

where:

V = velocity (ft/sec)
C = Hazen-Williams Coefficient
D = pipe diameter (ft)
S = slope
Q = flow (million gallons / day [mgd])
0.2795 = conversion factor from ft^3/sec to mgd.

If, in the actual system, the pipe actually has a gradient greater than 0.01 the actual gradient was used for S in the above equation. Maximum flow for pumps was determined by the maximum pump capacity, not by the above procedure. In practice this estimate worked very well, except for links representing links between tanks and the system. For these links it was finally easiest to determine the limits of flow by examining a few configurations as solved by a computer program which can solve for the flows and pressures in the hydraulic network (based on the algebraic equations of network flow). For these analyses, SDP8

(1984), a computer package for such hydraulic calculations by Charles Howard and Associates, Ltd. was employed. The rationale for the limits used for these links in each sample system will be discussed, as the system results are presented.

Analysis of Network A (Figure 3-6), proceeded very smoothly. The limits for most of the links were determined from formulas 3-2 and 3-3 above, with a slope of 0.01. Links 98 and 99, the pressure reducing valves, have a gradient steeper than 0.01. Thus for these links the actual slopes of 0.37 and 0.44 were used. Lastly, this system has only one pump, so the maximum capacity of this link was set to the total demand of the system. Table 3-6 presents the calculations and the link capacities used for this analysis. Table 3-7 contains the demand at each node.

Link	From	To	C	D (in)	S	Capacity (mgd)
1	2	3	120	16	0.01	5.946
2	3	4	120	12	0.01	2.790
3	3	6	120	14	0.01	4.185
4	4	5	120	10	0.01	1.727
5	6	5	120	14	0.01	4.185
6	8	7	120	8	0.01	0.960
7	8	9	120	10	0.01	1.727
8	7	9	120	8	0.01	0.960
9	10	7	120	10	0.01	1.727
10	7	10	120	6	0.01	0.451
98	4	10	65	6	0.37	1.716
99	5	8	65	4	0.44	0.649
100 (pump)	1	2	–	–	–	6.675

Table 3-6 Network A Link Capacity Data

Node	Demand (mgd)
1	-
2	1.6
3	1.2
4	0.6
5	0.4
6	0.875
7	0.6
8	0.8
9	0.4
10	0.2

Table 3-7 Demands for Network A

Calculation of the probability that every node receives sufficient supply for this series-parallel network required very little time for analysis, less than 3.0 CPU seconds on a VAX 11/750 per calculation. Thus a variety of runs were performed, both to throughly analyze the system and to investigate the suitability of the model.

Table 3-8 presents the probability of sufficient supply for a number of cases involving Network A. Connectivity values for each case are included for comparison. The first set of analyses were performed to investigate the sensitivity of the analysis to the probabilities of operation of each link. Thus the following cases, as used in section 3.2.2, were run:

* all links have $p_i = 0.95$
* all links have $p_i = 0.97$
* all links have $p_i = 0.99$

* all links have $p_i = 0.999$
* link 1-2 has $p_i = 0.9543$ (pump) and the rest have $p_i = 1 - (1.557 \times 10^{-6}) \times$ pipe length (ft).

Secondly, a number of capacity conditions were investigated. Runs were performed with the following conditions:

* normal capacity (Table 3-6)
* 0.9 x normal capacity
* 1.1 x normal capacity
* 1.2 x normal capacity.

For each of the above runs all link probabilities were set to the last case above, since the probability of operation of a link is expected to depend on whether a link is a pipe or a pump, and on the length of the pipe.

Case	Link Probabilities	Capacities	Probability of Sufficient Supply	Connectivity (from Table 3-2)
I	all = 0.95	Table 3-6	0.6586	0.8902
II	all = 0.97	Table 3-6	0.7817	0.9364
III	all = 0.99	Table 3-6	0.9225	0.9796
IV	all = 0.999	Table 3-6	0.9920	0.9980
V	Table 3-3	Table 3-6	0.9426	0.9540
VI	Table 3-3	1.1 x Table 3-6	0.9476	0.9540
VII	Table 3-3	1.2 x Table 3-6	0.9483	0.9540
VIII	Table 3-3	0.9 x Table 3-6	0.9426	0.9540

Table 3-8 Supply and Connectivity Probabilities for Network A

A comparison of the probabilities of sufficient supply with the corresponding connectivity values indicates, as expected, the probability of the system being able to deliver the required supply is less than the probability of it simply being connected. Particularly for the sets with less reliable links, the probability of sufficient supply is considerably less than the probability of being connected. The wide range of these probabilities is quite striking. For a system with $p_i = 0.95$, the probability of sufficient supply = 0.65866, compared with a probability of sufficient supply = 0.9920 for a system with $p_i = 0.999$.

For Network A, the supply probabilities also seem to be linearly related to the link probabilities, for cases with equal p_i's. (Figure 3-13) For the last, more realistic, probability set under normal capacity conditions, the probability of sufficient supply is 0.9426, which is fairly close to the connectivity of 0.9540. For systems with higher demands, more demand points, and/or pipes operating closer to capacity under normal operating conditions, the difference in these values would be expected to be larger.

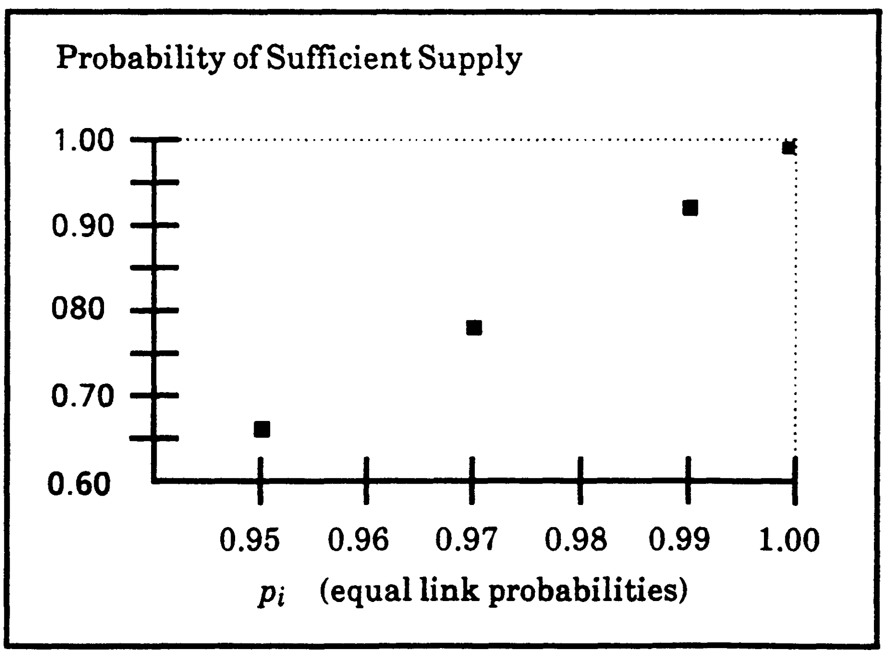


Figure 3-13 Graph of Probability of Sufficient Supply versus Link Probability for Network A

The 0.9 x normal capacity case had the same feasible minimum configurations, and the same probability of providing sufficient supply as the normal capacity case. It is interesting to examine the configurations "on the edge", namely those feasible only under capacity conditions greater than normal. Each such configuration is a "minimum link set" under normal capacity, implying if any one of the arcs shown fails the configuration is no longer feasible. Figure 3-14 presents all the configurations which were found to be feasible under for the normal capacity, followed by additional configurations which become feasible as the capacities increase.

For capacities in the range used for these examples, links 1,2,3, 98, and 100 must operate for the system to operate. Link 99 (5-8) is necessary for all but two configurations in the 1.2 x normal capacity case, indicating this this link is also quite crucial. For the normal capacity case *only*, the link 5 (4-5) is never in the minimum link set, indicating the loss of this arc never causes the system to fail. However, for some of the minimum feasible configurations for the 1.1 x normal capacity case, link 5 is included. Thus it is expected that link 5 will add a small amount of reliability to the system. Link 10 (10-7), which is parallel to another link, is never in the minimum link set for any capacity condition. Thus link 10, perhaps as well as link 5, might be removed without affecting the reliability of the system to a large extent.

Calculation of the probability of sufficient supply for Network B (Figure 3-10), was somewhat more difficult than for Network A. Network B, with 16 links, and sections such as the pentagon around node 90 has much more redundancy than network A, in the sense that many more alternative feasible configurations are

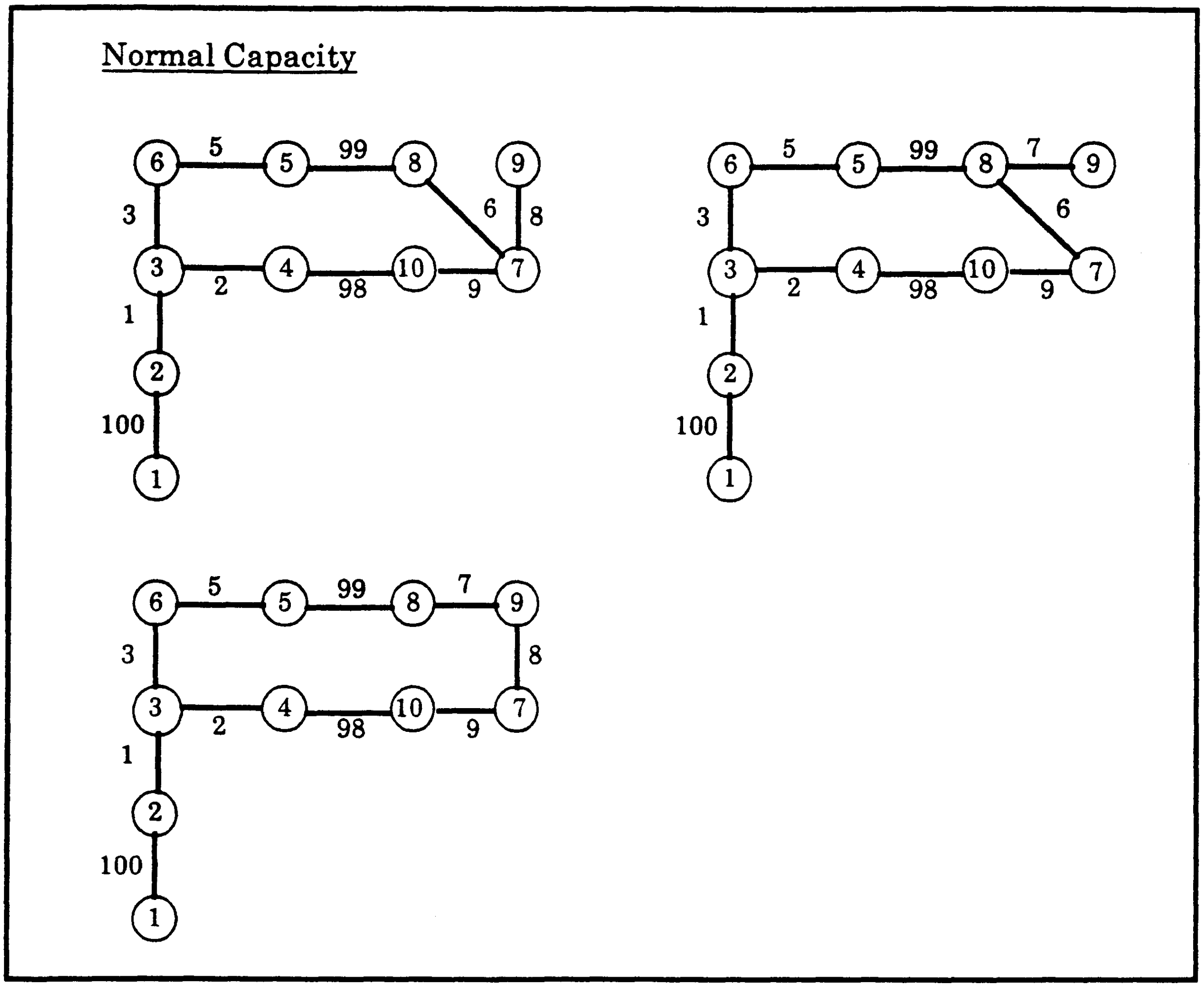


Figure 3-14 Supply Feasible Configurations for Network A

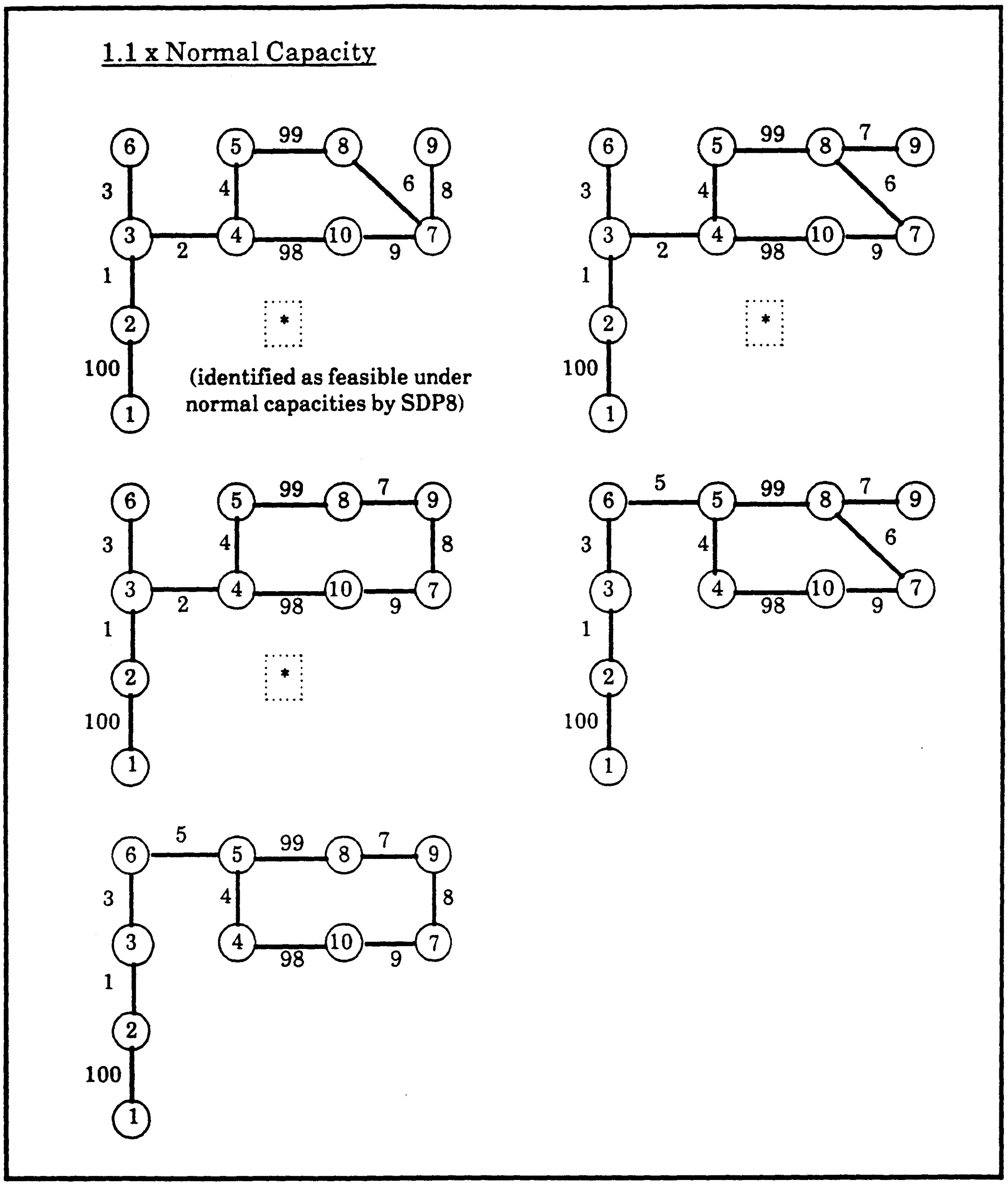


Figure 3-14 Supply Feasible Configurations for Network A (continued)

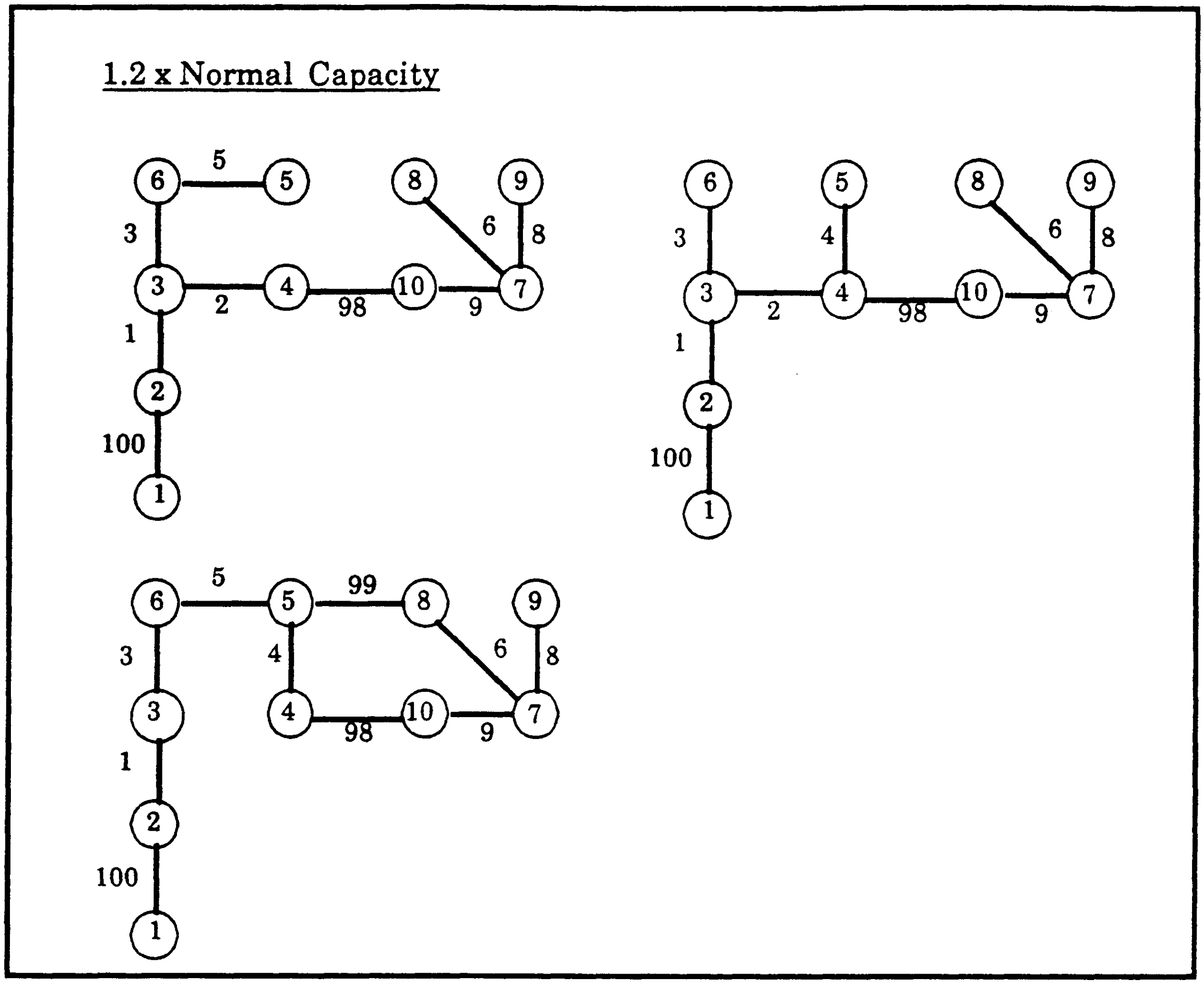


Figure 3-14 Supply Feasible Configurations for Network A (continued)

possible. Thus Lee's algorithm, which finds feasible configurations, took considerably longer to analyze the larger network.

Every pipe in Network B has a gradient of less than 0.01, so the pipe capacities can be determined from formulas 3-2 and 3-3, with S=0.01. However, for this system demands were specified in gallons / minute (gpm), so the values from equation 3-3 must be multiplied by 694.44 to convert from million gallons / day. For the capacities of the pumps and the links connecting the water tanks, it was necessary to investigate the behavior of the system with SDP8. Table 3-9 presents the calculations and the link capacities used for this analysis. Table 3-10 contains the demand at each node.

Link	From	To	C	D (in)	S	Capacity (gpm)
2	20	70	70	16	0.01	2408
6	20	110	70	12	0.01	1130
10	70	100	70	12	0.01	1130
12	70	90	70	10	0.01	700
14	60	70	70	12	0.01	1130
16	60	90	70	10	0.01	700
18	60	80	70	12	0.01	1130
20	80	90	70	10	0.01	700
22	90	150	70	10	0.01	700
24	90	100	70	10	0.01	700
26	100	150	70	12	0.01	1130
28	80	150	70	10	0.01	700
48	100	110	70	8	0.01	389
78 (tank link)	65	60	-	-	-	1500
101 (pump)	10	20	-	-	-	3000
102 (pump)	10	20	-	-	-	3000

Table 3-9 Network B Link Capacity Data

SDP8 results show the tank must supply water, and one pump must be operating for the system to meet demand at each node. When one pump is out the tank supplies about 1500 gpm, so the capacity of the tank connection was set to 1500 gpm. By similar reasoning, each pump capacity was set to 3000 gpm, to allow the 4200 gpm demand to be met with one pump missing. Although these limits on the pumps and tanks link capacities seem to allow the system to operate without the

Node	Demand (gpm)
10 (river)	-
20	500
60	500
70	500
80	500
90	1000
100	500
110	500
150	200
65 (tank)	-

Table 3-10 Demands For Network B

tank suppling water, the capacity limits on the pipes near the tank do not permit the network to operate without the tank in operation.

Calculation of the probability of sufficient supply for this non-series-parallel network required approximately 2 CPU minutes on a VAX 11/750 per calculation. Thus a smaller number of link probabilities and capacity cases were performed than for Network A. Table 3-11 presents the probability of sufficient supply for a number of cases involving Network B. The corresponding connectivity values are included for comparison. Only four calculations were performed to investigate the sensitivity of the analysis to the probabilities of operation of each link, as follows:

* all links have $p_i = 0.95$
* all links have $p_i = 0.97$
* all links have $p_i = .99$
* links from 10-20 have $p_i = 0.9543$ (pumps) and the rest have $p_i = 1 - (1.577 \times 10^{-6}) \times$ pipe length (ft) [Table 3-5].

Secondly, only the following three demand conditions were investigated:

* normal demand (Table 3-10),
* 0.9 x normal demand
* 1.1 x normal demand.

For each of the above runs the probability of operation of each links were set to the last probability case.

Case	Link Probabilities	Capacities	Probability of Sufficient Supply	Connectivity (from Table 3-5)
I	all = 0.95	Table 3-9	0.8463	0.8956
II	all = 0.97	Table 3-9	0.9085	0.9383
III	all = 0.99	Table 3-9	0.9698	0.9798
IV	Table 3-5	Table 3-9	0.9514	0.9534
V	Table 3-5	0.9 x Table 3-9	0.9511	0.9534
VI	Table 3-5	1.1 x Table 3-9	0.9514	0.9534

Table 3-11 Supply and Connectivity Probabilities for Network B

Again, a comparison of the probabilities of sufficient supply with the corresponding connectivity values indicates, as expected, the probability of the system being able to deliver the required supply is less than the probability of it simply being connected. The range of the probabilities is less than that of Network A, with a probability of meeting the specified supply of 0.8463 when all links have $p_i = 0.95$, as compared with 0.9798 when all links have $p_i = 0.99$.

Network B has many more minimum link feasible configurations than Network A, three for Network A versus 289 for Network B. Again for this network, the supply probabilities also seem to be linearly related to the link probabilities, for cases with equal p_i's (Figure 3-15). For the last, more realistic, probability set under normal capacity conditions, the probability of sufficient supply is 0.9514, which is quite close to the connectivity of 0.9534.

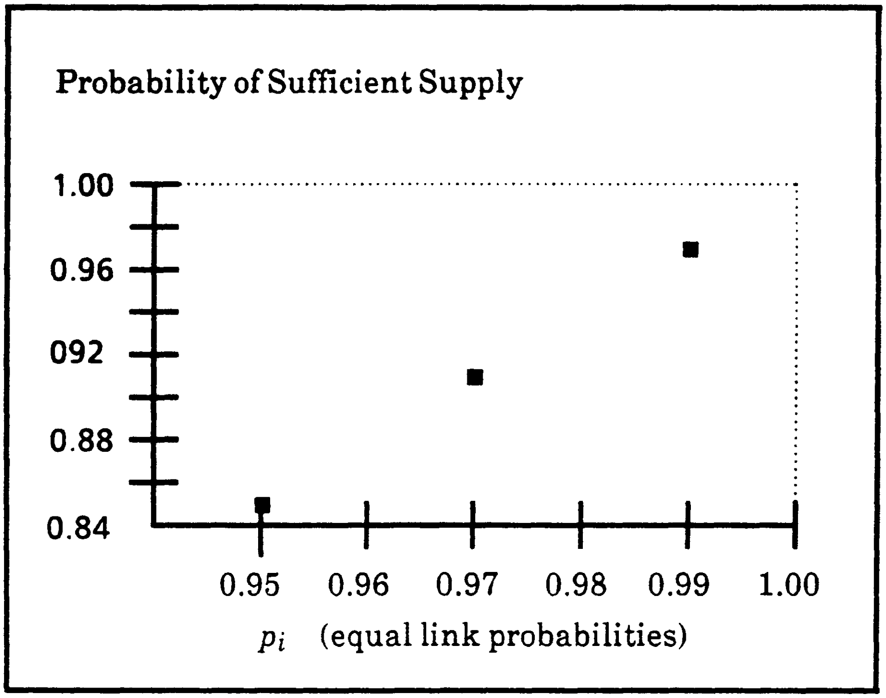


Figure 3-15 Graph of Probability of Sufficient Supply versus Link Probability for Network B

In all of the feasible configurations identified, the tank link (78) and at least one pump were operating as required. Also, links 2 and 6 (the links leading out of node 20) were operational in all feasible capacitated network supply solutions identified. Most of the configurations involved different combinations of failed links from the pentagon around node 90. Thirty-five feasible configurations were identified with 5 failed links in the pentagon. Five links failed in the pentagon is the maximum possible number of failed links; six or more link failures in the pentagon cause some node to become disconnected. For all of these 35 configurations, links 2, 6, 10, 78, and at least one pump operate. Link 14 never operates in any of these 35 configurations, although it does operate in some of the other 254 identified feasible configurations. Eight configurations were identified with five links out in the pentagon, plus one pump and link 48 (from node 100 to 110) failed. These eight configurations are in fact "spanning trees", *e.g.* a network with no cycles containing all the nodes. The eight spanning tree solutions are shown in Figure 3-16.

These spanning tree solutions are the best candidates for "on the edge" configurations. Increasing or decreasing the link capacities by 10% did not appreciably change this reliability measure for this network.

3.3.3 DISCUSSION

As demonstrated by some of the previous examples, the probability that a network can provide sufficient supply at each node may be considerably less than the probability that it is simply connected. Thus except for systems in which every pipe carries supply well under its capacity the supply based reliability measures should be used.

The network model used for these calculations suffers from the same problems as the reachability and connectivity measures in the use of p_i values to represent the reliability of each link. Again the uncertainty introduced by the p_i's will be very small compared to the uncertainty in the data used to estimate these values. Thus the use of these instantaneous probabilities, instead of ones accounting explicitly for the operate-fail-operate cycles should not to be of concern.

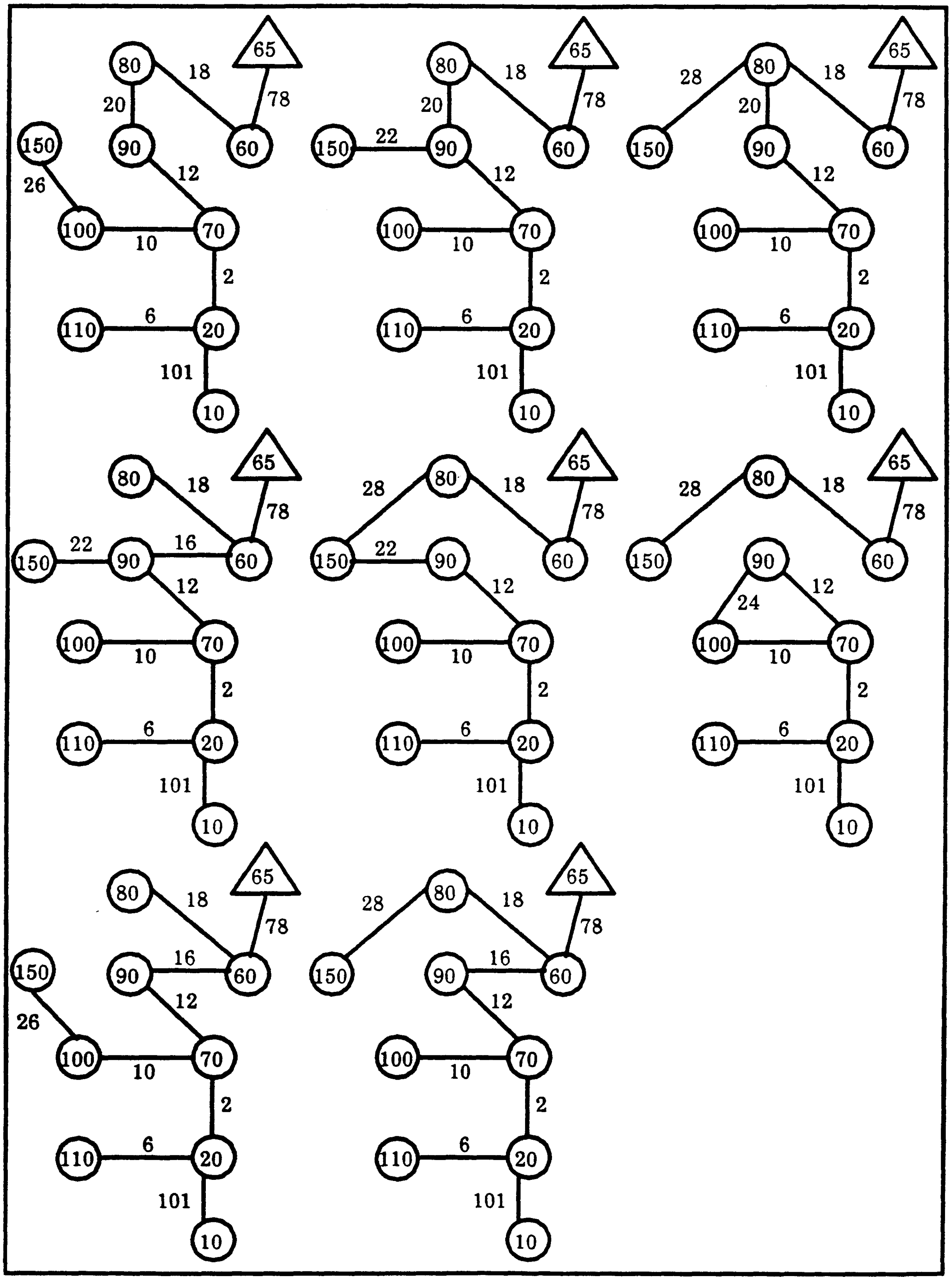


Figure 3-16 Feasible Spanning Trees for Network B

Of more concern, is the use of explicit limits on pipe flow capacities. As previously discussed, pipes in networks do not truly have maximum capacities, since an increase in pump capacity can always force more water through the pipe (as long as the pipe does not burst). Additional tests were run with Networks A and B in an attempt to determine how well the water distribution systems were represented by capacitated network models. This further analysis consisted of a check of the configurations identified as providing sufficient supply in the capacitated networks with SDP8. With SDP8, fairly large networks can be solved for either supplies or pressures at each demand node, for a network in which pipe lengths, resistances, pump characteristics, and either demand or pressure at each node are known. Of course, using SDP8 requires more computational effort than finding the minimum cost flows to meet supply on a capacitated network.

SDP8 was used to model the networks by specifying:

* the length and resistance of each pipe,
* the demand at each demand node, and
* the pressure at each supply node.

Thus the unknowns, calculated by SDP8, were the supply at each supply node, and the pressure at each demand node. A configuration was said to actually be feasible if the pressure, as calculated by SDP8, for every node was at least 40 pounds / in^2 (psi). (In Chapter 4, different, more complicated definition of feasible based on pressures at each node was used. However such complexity is not appropriate for this simple model.)

For Network A there were few enough configurations so that the probability of sufficient supply could be calculated with SDP8 "linked in" to Lee's algorithm. All of the configurations previously identified as feasible when the normal link capacities were used were also feasible using SDP8. There were however, three additional feasible configurations identified only by SDP8. These configurations are in fact the first three configurations identified as feasible in the 1.1 x capacity case (Starred in Figure 3-14). All of these configuration were very close to the 40 psi limit, with the minimum pressure (always at Node 9) at 42 psi or less. The additional reliability accounted for by these configurations is less than 0.005.

The two other configurations identified as feasible in Network A for the 1.1 x capacity case, were also tested with SDP8. Both had minimum pressures in the range of 20 - 25 psi. The 1.2 x capacity case allows three additional configurations, however, two of these when checked with SDP8 were not feasible (had negative pressures at some nodes).

As an additional test of Network A, the probability of sufficient supply was recalculated with SDP8 with the definition of an operating system changed to: all nodes must have pressure of at least 20 psi. Four additional configurations, with minimum pressures in the range of 20-30 psi were identified.

For Network A at least, with the procedure described for calculating pipe capacities, this method does come very close to calculating the correct probability of sufficient supply. Also, by varying the pipe capacities by small amounts, useful information about the contributions made by individual components to the overall reliability can be gained.

Network B was not so easily checked with SDP8. Recalculating the probability of sufficient flow for this network with Lee's algorithm linked to SDP8 is not practical, due to the computer time and space requirements of such an analysis. However, the eight spanning trees identified in Figure 3-16 were analyzed with SDP8. Unfortunately, all eight had pressures at one node (150) below 40 psi.

However, only node 150 was below 40 psi for any of these configurations; the calculated pressures at node 150 for these configurations ranged from 16.4 to 36.4 psi.

A few other spanning tree solutions, not identified as feasible for the capacitated network, were examined by SDP8. Each of these "extra" configurations had 5 links failed in the pentagon, but never links 10, 12, or 18. These configurations also had minimum pressures at node 150, in the range of 28.7 - 34.9 psi. Thus the capacitated network model is perhaps somewhat imprecise at identifying feasible configurations "on the edge".

In all of the eight spanning tree configurations, links 10, 18, and 12 were operating. A few configurations with some of these links failed were also analyzed with SDP8. Indeed all of these configurations had nodes with pressures well below 40 psi (most had nodes with calculated negative pressures). Thus the capacitated network solutions appear, from this limited sample, to distinguish clearly infeasible configurations.

Note, for Network A the configurations not found by the capacitated network solution, or found incorrectly, all involve a number of failed links. This same pattern appears to be true for Network B as well. For highly reliable components, the probability of having a number of links out at the same time is very small. Thus this method is likely to be more accurate for networks with fairly high p_i's.

In conclusion, the use of Lee's algorithm and a model of capacitated network flow appears to hold promise as a analytical method for the calculation of the probability of sufficient supply in a network. This reliability measure was checked with SDP8 for Network A; the probabilities from the two models were very close. Network B could not be checked this way, but the investigations on selected configuration with SDP8 appear to support the applicability of this method. Additional information on the accuracy of these probability values can be found by simulation, the topic of the next chapter.

4.0 SIMULATION METHODS

4.1 USE OF SIMULATION FOR RELIABILITY CALCULATIONS

As shown in the previous chapter, certain reliability measures defined on general networks can be calculated analytically. These analytical methods can provide a fast initial assessment of the reliability of a system. However, all of the analytical methods developed thus far involve fairly stringent assumptions which can limit the applicability of analytical results to understanding real-life systems. Also the measures that can be calculated analytically are limited. Small systems with features such as water storage tanks and operational responses to failures within the system have been investigated analytically with fault-tree analysis methods, however many of these methods also involve a number of simplifying assumptions and approximations. Thus for a thorough investigation of the reliability of a water distribution system, stochastic simulation methods incorporating more complicated and realistic features of the system may be desirable. In particular, for moderately large water distribution networks (10 – 50 nodes), simulation methods can provide accurate estimates of reliability for systems with elements not easily incorporated in analytical methods. Thus, once an initial assessment of the reliability of a water distribution system has been performed analytically and alternative improvement options proposed, a simulation of these options should be done to gain a better understanding of how the proposed alternative systems will be likely to behave under real-life conditions.

This chapter presents an event-oriented, discrete simulation program developed to assess the reliability of water distribution networks subject to failure due to pipe breaks and pump outages. This program can be used to calculate a variety of reliability measures relating to the number, location, duration, and effects of failures. This simulation approach allows great flexibility in the types of network elements that can be included for analysis, and of the failure time and repair time distributions. Additionally, changes in the operation of the system in response to pipe and pump failures can be simulated. Evaluation of options for improving the system reliability can also be performed with this program.

In this chapter, the scope of the simulation is discussed in section 4.2. Two systems (plus a subset of one system) were simulated as examples of the use of this simulation program; the two systems are described in section 4.3. Following is section 4.4 which specifies the failure and repair time probability distributions used in this simulation. Section 4.5 contains some of the programming and statistical details of this analysis. Results of the simulations are given in Section 4.6. Section 4.7 discusses the applicability and usefulness of the simulation approach.

4.2 SCOPE OF SIMULATION

As in the previous chapter, the major source of unreliability in this analysis is failures due to random pipe breaks and pump outages. In contrast to the analytical methods of Chapter 3 however, this simulation analysis could be easily expanded to include sources of unreliability such as lack of supply (*e.g.* drought), and random contamination. Currently, only reservoirs, pumps, pipelines, and storage tanks will be allowed as components of the water distribution system.

The program is event-oriented, meaning the time step of each iteration is calculated as the time between consecutive random events. The time step then varies from iteration to iteration. The program consists of two parts:

1) the simulation section, which generates failure and repair events according to specified probability distributions, and

2) the hydraulic network solution section, which gives the flows throughout the network and the heads at each node, for a specified demand and system configuration.

The program SDP8 (1984), by Charles Howard and Associates, Ltd. which was used and described in Chapter 3 was incorporated as a subroutine in this simulation program.

The simulation program is designed to analyze fairly detailed water distribution systems, but is not appropriate for modeling every connection to every house in the city. Instead, it is assumed demand for water is aggregated into demand points, which are represented by nodes in the water distribution network. Connecting pipes, as well as pumps, are represented by links between these nodes.

For this study, demand is assumed to be known and constant. Thus the fluctuations of demand over the day and over the year, though expected for any real system, will in this study be "smoothed" into one average daily demand. Since much of the data about failures is not well known, any reliability analysis is approximate. Thus it is assumed the additional approximation of using an average specified demand is appropriate. With a similar rationale, it will be assumed at the initiation of each set of failure events, the water storage tanks will be filled to their average levels.

Since SDP8 is used as a subroutine, the simulation program was designed to use the SDP8 data input file. To start the simulation the following information must be specified:

* configuration of the network
* demand (in terms of volume/time) at each demand node,

* head at each supply node,
* Hazen-Williams coefficient for each pipe,
* pump curve for each pump,
* geometry of the water tanks,
* estimated head at each demand node under normal conditions, and
* estimated supply (in terms of volume/time) at each supply node under normal conditions.

In SDP8 the last two sets of estimated data are calculated exactly through an iterative procedure. These estimated values are used as a starting point for the SDP8 solution. Thus to reduce the computations necessary for the simulation program, the full network should be preproccessed by:

(1) running SDP8 for the full system with data (both known and estimated) as specified above,

(2) replacing the estimated values with the values calculated by SDP8.

This preprocessing ensures that SDP8 will converge quickly when it is used as a subroutine to analyze the system with only a pipe or two failed.

The forms of the probability distributions used for pipe, and pump, break and repair times are fixed by the simulation program (see Section 4-4). The parameters for these distributions however, are user-specified and must be input in an auxiliary data file. The simulation proceeds by randomly generating failure times of the pipes and pumps according to the specified failure time distributions. When a link fails it is removed from the system. The new heads at the demand nodes in the reduced network are determined by generating a new SDP8 input file, and then making a call to subroutine SDP8. It is assumed link failures leave the demands unchanged. The new heads at the demand nodes are used to tell how the system is performing.

In this simulation it is assumed the water supplied to a given node will depend on the head attainable at the node. For each node, two head limits must be given:

(1) a minimum head (H_m), and
(2) a service head (H_s).

These head limits are also stored in the auxiliary data file.

The system will be said to be performing normally only when, for each node, all the imposed demands can be met with heads above the service limit. If however, at some node in the diminished system, the head is below the service limit, it is assumed at that node the system cannot supply the full demand. In this case, it is assumed the operators of the system react with the following policy:

(1) nodes with heads below the minimum head limit will be completely shut off,
(2) nodes with heads above the minimum head limit but below the critical head limit will be supplied at a reduced level, and
(3) nodes with heads above the service level are supplied at normal service levels.

Each node can then be in a normal, reduced service, or failure mode. (See Table 4-1). The system will be said to be in normal mode if *all* nodes are receiving normal supply, in failure mode if supply to *any* node has been shut off, and in reduced service mode if *some* node or nodes are receiving reduced supply but *no* nodes are completely shut off.

As a simple but general purpose policy, supply for a node in reduced mode will be reduced according to the following equation:

$$Q = [(H - H_s) / (H_m - H_s)]^{1/2} * C \qquad (4\text{-}1)$$

Mode	Head at Node (H)	Supply
Normal	$H_{service} < H$	Equal to demand
Reduced Service	$H_{minimum} \leqq H \leqq H_{service}$	Reduced according to equation 4-1
Failure	$H < H_{minimum}$	None

Table 4-1 Node Service Modes

where:

H_s = service head,
H_m = minimum head,
H = calculated head ($Hs \leqq H \leqq H_m$), and
C = full demand at the node.

Figure 4-1 shows how the supplied flow varies with head for a typical node. The rationale for this formula is that hydraulic laws for flows through devices show flow is proportional to the square root of head. Thus we assume the supply reduction from normal supply to no flow will be related to the square root of the computed head level.

Once a link has failed, a random repair time is generated and the system is assumed to operate in the diminished state until the repair time is reached, or another link fails. The simulation program records the total duration of reduced service and failure periods at each node, the shortfall at each node, and various information as listed in Table 4-2.

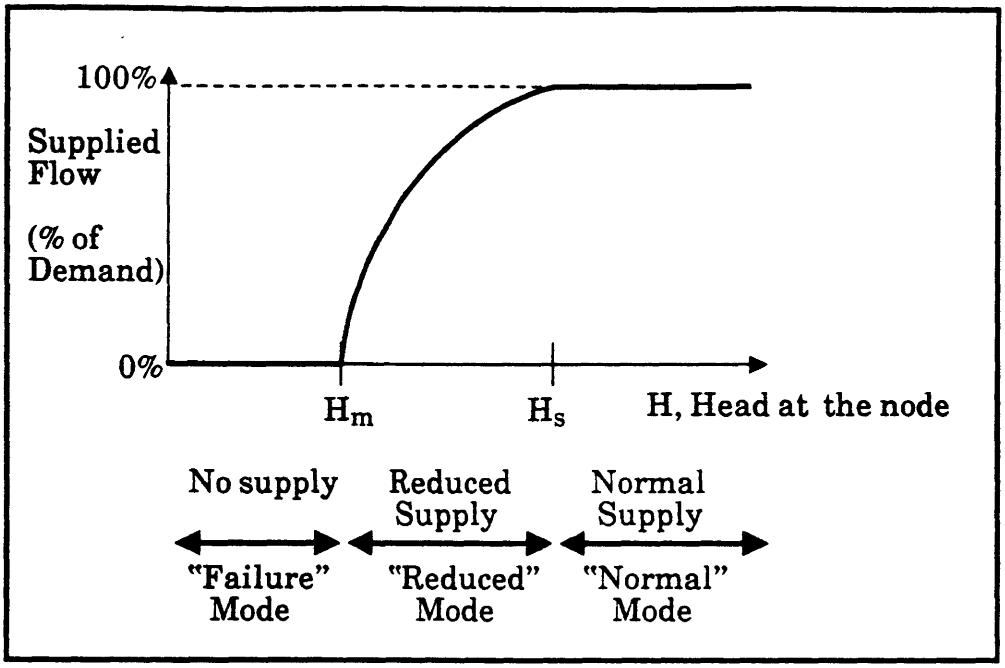


Figure 4-1 Supply Response to Head Loss

4.3 DESCRIPTION OF SIMULATED SAMPLE SYSTEMS

Two sample systems were selected to test and demonstrate this simulation method: 1) a small 10 node network and 2) a larger 22 node network. These systems were chosen because they provide a good range of features to test the program. The two systems will be referred to by the innovative names of System A and System B+.

System A (Figure 3-6), which was also analyzed in Chapter 3, is a small series-parallel network with 9 demand nodes, 1 reservoir, 10 pipes, 1 pump, and 2 pressure reducing valves. Table 4-3 presents the elevations of the nodes in System A. Note, in this system water must be pumped up from the reservoir and the pumping station (nodes 1 and 2 at 100 feet) to a higher elevation zone (nodes 3, 4, 5, and 6 at 200-350 feet), and then down again to a zone of lower elevation

Event Related:
Type of event (failure or repair) Duration of event Total number of events in simulation period System status during each event (normal, reduced service or failure)
Node Related:
Total demand during simulation period Total supply during simulation period Shortfall Average head Minimum head during non-failure periods Minimum consumption during non-failure periods Number of reduced service events Duration of reduced service events Number of failure events Duration of failure events
Link Related:
Number of pipe failures Total duration of failure time for each pipe Percentage of failure time for each pipe Number of pump failures Total duration of failure time for each pump Percentage of failure time for each pump
System Related:
Total system consumption Total number of breaks Time-averaged number of breaks Maximum number of breaks

Table 4-2 Accounting Information Kept for Each Simulation Run

(nodes 7, 8, 9, and 10 at 10-50 feet). Thus we expect the pump will be crucial to the reliability of this network, and reliability at the higher nodes will be less than at the lower ones.

Node	Elevation(ft)
1 (Resevoir)	100
2	100
3	200
4	210
5	230
6	350
7	10
8	10
9	50
10	25

Table 4-3 System A Elevations

The second system, System B+, (Figure 4-2) is a larger and more realistic network, with 16 demand nodes, 3 pumps, 2 water tanks, 1 water source, and 34 pipes. It is more highly connected, and not series-parallel. This network is taken from a water network problem proposed by Tom Walski, Army Corps of Engineers, as a design problem to be worked on by various water distribution network designers and presented at the July 1985, ASCE, Water Division Conference. Walski posed the problem as finding an economical set of older pipes to re-line or replace, and a few possible new pipes to install, so as to meet future demands on the system. As a first pass at a reliability analysis of this system, the current (1985) system under the current "average" water demands will be analyzed in this chapter. Table 4-4 presents the elevations of the nodes in System B+. As does System A, System B+ involves pumping water uphill. The water source is at a low elevation (node 10 at 10 feet), so water is pumped uphill from the river to downtown (nodes near 70, at 50 feet), to the new part of town (node 140, at 80 feet), and to the other town areas (nodes near 160, at 120 feet).

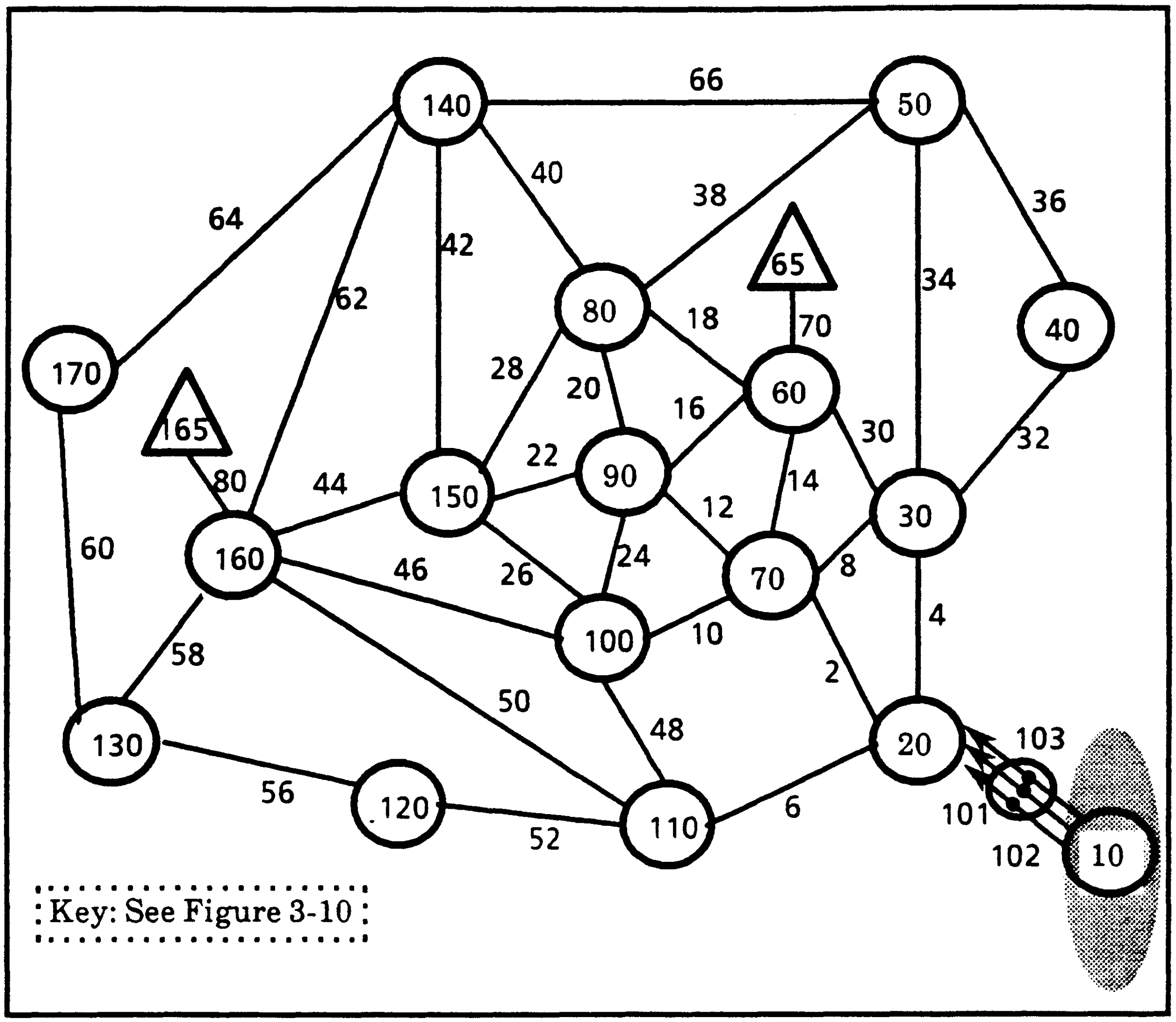


Figure 4-2 Network B+

As a preliminary investigation of the reliability of Network B+, several SDP8 runs were performed on different system configurations. The following configurations were examined:

* only one pump operating, both tanks operating,
* all pumps operating, both tanks failed (empty).

Node	Elevation (ft)
10 (river)	10 (average)
20	20
30	50
40	50
50	50
60	50
65 (tank)	235 (average)
70	50
80	50
90	50
100	50
110	50
120	120
130	120
140	80
150	120
160	120
165 (tank)	235 (average)
170	120

Table 4-4 System B+ Elevations

Even with only one pump, at every node the demand can be met at a head greater than or equal to the service head limit. However, when both tanks have failed every node is below the minimum head limit, implying when both water tanks are empty none of the demands can be met. Thus, the pumps are not likely to be as crucial to the reliability of this system as they were to the smaller system. However, the volume and operation of the water tanks is expected to have a large effect upon the reliability of the entire system, since the water tanks are so crucial to meeting average demands.

The second system analyzed in Chapter 3, System B (Figure 3-10), is a subset of Network B+. The nodes in Network B correspond to the part of the Network B+, described as being built before 1940. One simulation run for Network B was performed as a check on the analytical results.

4.4 FAILURE AND REPAIR TIME PROBABILITY DISTRIBUTIONS

Data about probability distributions of failure times and repair times for pipes and pumps are usually not readily available. No failure or repair time data specific to the two networks analyzed exist. Thus, for this simulation, "reasonable" distributions with "reasonable" parameters were chosen. The distributions and the parameters used in this simulation, as well as literature supporting these choices are described in this section.

A few studies have looked at quantifying the number of pipe breaks per unit time per pipe length, based on pipe qualities such as age, material, etc. Walski [1984b] and O'Day [1982] present some data on pipe break interarrival times but only look qualitatively at factors affecting these interarrival times. Shamir and Howard [1979] developed an exponential model describing the increase of pipe breaks with pipe age. Walski and Pelliccia [1982] added corrections to this model for the factors of pipe size and number of previous breaks. However, neither of these models looks specifically at the interarrival time between individual breaks of the same pipe. A study in the mid-seventies for the City of Ann Arbor, Michigan (reported by Pollock, personal communication 1985) found break rates for pipes were constant once the pipes were about 50 years old and interarrival rates for pipe breaks were exponential. It is commonly assumed, by these and other studies, that failures of different pipes occur independently.

For this study an exponential distribution was chosen for the pipe break distribution. A listing of the rate of pipe breaks for various US cities, obtained from a US General Accounting Office report to Congress [1980], is presented in Table 4-5. A figure of 1 break/1 mile/year was picked for use for all pipes. This figure is in the high range so as to fully exercise the simulation system. For each pipe this is multiplied by it's length to give the average number of breaks per year for the link. The reciprocal of this number is the exponential parameter - the expected time between breaks. Figure 4-3a presents this distribution graphically for a 0.5 mile long pipe.

Marks, *et. al.* [1985] presents a hazard failure model giving the probability, at any dt, that the pipe will break, based on several factors including the age of the pipe, the number of previous breaks, and the time since the last break. This hazard function could be added to this simulation program as an arrival rate for a non-homogeneous Poisson process. For this first study, such an analysis seemed overly complicated.

There is even less data available on pump breaks than on pipe breaks. In a simulation of a water distribution system with only pump failures Damelin, *et . al.* [1972] used an exponential distribution for pump break interarrival times. However, their data were based on interarrival times of working hours, thus ignoring times when the pumps were inoperative due to scheduled outages for maintenance. A representative time of 1000 hours was chosen as the mean time between pump failures from their paper. This pump break time distribution is graphed in Figure 4-3b.

City	Year	Pipe breaks/ 1000 miles/ year
Boston	1969-70	36
Chicago	1973	54
Denver	1973	156
Houston	1973	1290
Indianapolis	1969-78	83
Los Angeles	1973-74	43
Louisville	1964-76	123
Milwakee	1973	234
New Orleans	1969-78	680
New York City	1976	75
San Francisco	1973	106
St. Louis	1973	106
Troy, NY	1969-78	167
Washington, DC	1969-78	116

Table 4-5 Pipe Break Data

Source: "Additional Federal Aid for Urban Water Distribution Systems Should Wait Until Needs are Clearly Established, Report to the Congress", US General Accounting Office, Washington, DC, November 24, 1980; as quoted by Kelly O'Day, "Organizing and analyzing leak and break data for making main replacement decisions", p. 591, Journal American Water Works Association, November, 1982.

Again from Damelin *et. al.*, pump repair times can be represented by a log-normal distribution [Arad, 1968]. Parameters were chosen for the model based on the 50 hours mean time between failure given in Damelin but with more variability, again so as to fully exercise the program. Figure 4-3d presents the distribution used, ($\upsilon = 3.93$, $\sigma = 0.20$).

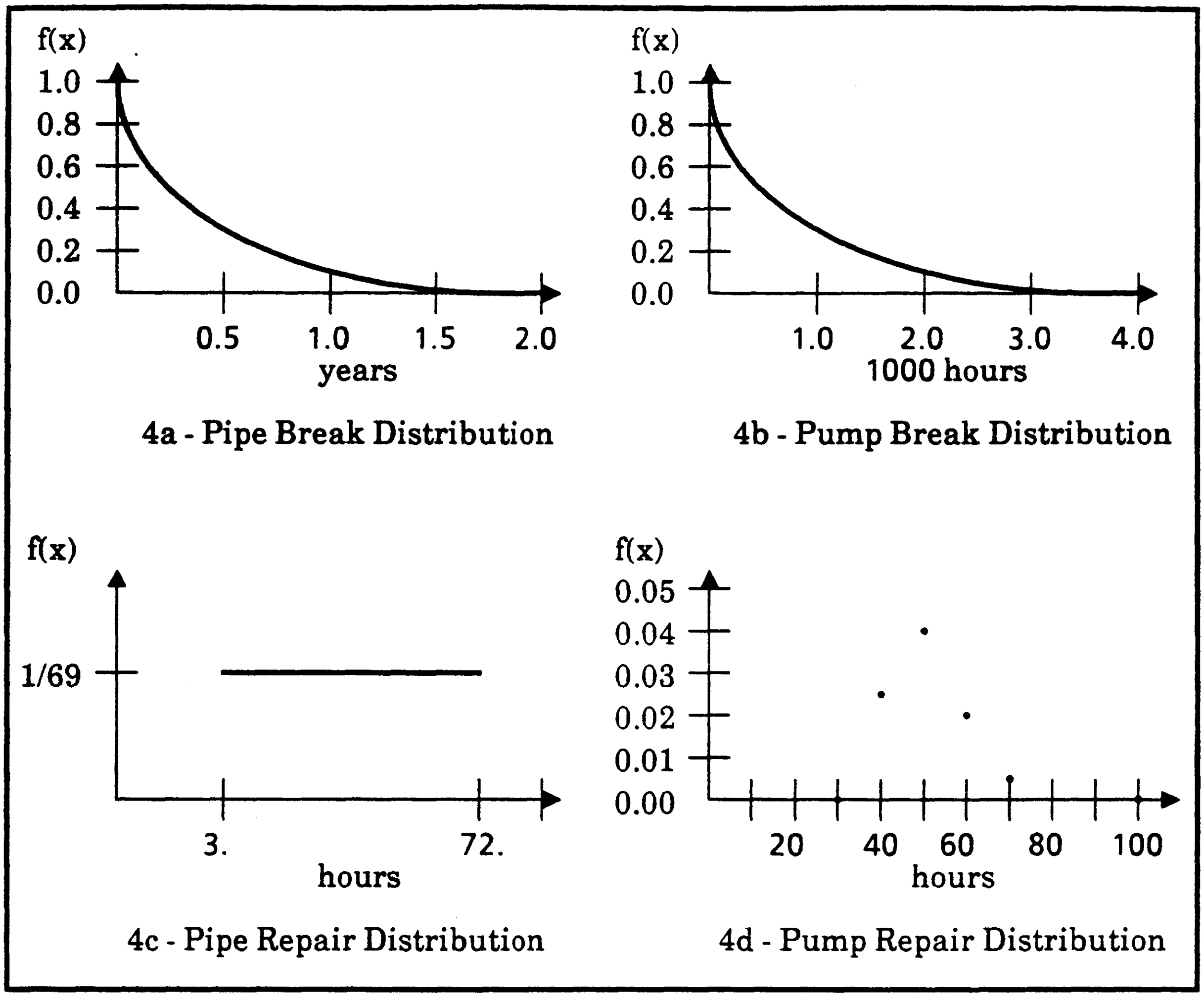


Figure 4-3 Break and Repair Time Probability Distributions

No data were found for pipe repair time distributions. A uniform distribution between 3 hours and 3 days (72 hours) was chosen as a first estimation of this process. (Figure 4-3c). For an analysis of an actual system, data on these times should be available from maintenance and payroll records of many urban public works departments.

4.5PROGRAMMING AND STATISTICAL DETAILS

As shown in Table 4-2 a number of reliability measures are calculated by this simulation program. Since these measures are based on a finite number of random events the calculated values are only approximate. To obtain some idea of the confidence intervals associated with these estimated values, a few reliability measures were chosen. These were annual shortfall (gallons), the percentage of time spent in emergency mode (for every node), and the percentage of time spent in failure mode (for every node). Although this choice was made to avoid the computational effort necessary to calculate confidence intervals for every criteria in Table 4-2, these measures are also felt to be some of the most important measures for assessing the reliability of the system. The shortfall measure is felt to be a good overall indicator of the reliability of the system. However, a low shortfall for some systems could be obtained by disconnecting, at any sign of emergency, one node of moderate demand so as to supply the others. Thus the percentages of time spent in the non-normal conditions for each node were also examined in evaluating alternative systems, to check for such "scapegoating". Examination of these measures also allows the identification of nodes markedly more or less reliable than average.

Since simulation results are not exact and will vary over (independent) simulation runs, comparisons of different systems must not be made on the basis of one simulation run for each alternative. Originally for this study 40 individual simulations of 5 simulated years each were planned for each system examined.

It is important to note how the statistics of, for example, annual shortfall as calculated from 40 runs of five years each compare with 1 run of two hundred years. First, in the 40 five-year simulations the system starts with all elements working. In the 1 two-hundred-year simulation there is no guarantee the system will be working on the first day of every fifth year. However, since both of these systems are fully operational much more than they are not, the disagreements due to this factor will probably be minor. Thus if we could perform two studies:

(1) an infinite number of 5 year runs, and (2) a simulation for an infinite time period, then the average annual shortfall calculated by both methods will closely agree.

The main reason for running 40 five-year simulations versus 1 two-hundred-year simulation is to be able to quantify the uncertainty in the reliability estimates. For example, a single average annual shortfall estimate (such as in the 1 two-hundred-year run) is a number with no variance. However, when the 40 five-year runs are used to estimate this value, and the estimated annual shortfall *and* the variance of this estimate can be calculated.

One way of calculating means and confidence intervals comes from classical statistics, by noting that most of the estimates involve the sum of a large number of random occurrences. Thus by a law of large numbers the reliability estimates are expected to be normally distributed. Then, for n estimates: $x_1, x_2, \ldots, x_n$, the mean X is estimated by:

$$* \; X = \left[\sum_{i=1}^{n} x_i \right] / n. \tag{4-2}$$

The variance σ is estimated by:

$$* \; \sigma^2 = \left[\sum_{i=1}^{n} (x_i - X)^2 \right] / n. \tag{4-3}$$

The 95% confidence interval (CI) is given by:

$$* \; CI = X \pm 1.96 \cdot (\sigma^2 / n)^{1/2} \tag{4-4}$$

For System A, a five-year simulation period was adequate and implementable. Five years is long enough for a number of elements (about 50) to break. Initially for System A, 40 runs of five years each were performed giving confidence intervals within 10% for shortfalls and for the percentage of time in failure mode for each node. Almost no time was spent in reduced service mode by any node, so these percentages were extremely small. The confidence intervals for such small numbers have little meaning. The overall conclusion, that time in reduced service mode is negligible, is nonetheless still valid. Thus 40 runs five-year simulations were judged to give sufficient accuracy for this problem. However, System B+ was not so easy to analyze.

System B+ is much larger than System A and many more breaks occurred in the same simulated time period. Initially, a five-year simulation period was tried, but this experiment resulted in over 500 breaks, exceeding the storage capacity of the program. As another experiment, a one-year simulation period was tried. However, the simulation of 40 runs of System B, even for only a 1-year period, turned out to be impractical in terms of required computer time. Thus a different approach to calculating confidence intervals, the regenerative method, was implemented for the System B+ simulations. (See Law and Kelton [1982, pp. 297 - 302] for a detailed description of the regenerative method.) The regenerative method involves measurements within a "cycle", which in this case can be defined as the time between successive times when the system first becomes fully operational. Within each individual cycle the shortfall, time spent in each mode, etc. *for that cycle* are tabulated. The n cycle times and measurements are statistically independent, and so can be used with methods from classical statistics to calculate the required reliability estimates and the associated confidence intervals. Thus for example, after n cycles, n cycle times $(c_1, c_2, \ldots, c_n)$ and n shortfall measurements $(s_1, s_2, \ldots, s_n)$ are obtained . The best estimate for the time-average shortfall S is given by:

$$S = \sum_{i=1}^{n} s_i \Big/ \sum_{i=1}^{n} c_i \qquad (4\text{-}5)$$

We can calculate an estimated annual average shortfall S_i for each cycle i by (note c_i in hours):

$$S_i = \left(s_i - S \cdot c_i\right) \cdot 8760 \tag{4-6}$$

The variance σ^2 of these cycles estimates is given by:

$$\sigma^2 = \left[\sum_{i=1}^{n} (S_i - S)^2 \right] / n \tag{4-7}$$

The 95% confidence interval is then given by:

$$CI = S \pm \left[1.96 \cdot (\sigma^2 / n)^{1/2} \right] / \left[\left(\sum_{i=1}^{n} c_i \right) / n \right] \tag{4-8}$$

Once the regenerative method was implemented, the simulations for System A were rerun. The results from the two methods, for both the estimation and the confidence intervals, were almost identical.

Early simulation runs, even for System A, turned out to take considerable computer time. This large computational requirement occurred because SDP8 was being called for each failure, even though many of the same elements (the pumps) were failing again and again. To reduce these repetitive calculations, another preprocessing step was added. Before the simulation was run, each link in turn was deleted, while all the other links remained in place. A SDP8 solution was obtained for each of the above system configurations and the calculated head, status, and (possibly reduced) supply at each node stored on a binary file. The simulation program was then changed slightly to read the binary file whenever the system contained only one failed element, instead of calling SDP8 again. Since repair times are, in general, much shorter than failure times it is rare for two elements to be broken at the same time. Thus the stored heads could be used for most failure events, and this programming modification reduced the necessary computer time considerably.

A simplified flowchart of the simulation run is presented in Figure 4-4. The program was verified by careful handchecks on truncated runs of 10 failures, using a table of random numbers generated with the same seeds as used in the simulation. The random number routine, supplied with the VAX FORTRAN 5 product was used, untested, for these simulations. The program was checked to ensure:

(1) the correct failure and repair times were being generated,
(2) program control was working correctly,
(3) the correct SDP8 input files were being generated,
(4) SDP8 was working as a subroutine exactly as it had when it was a stand-alone program, and
(5) the accounting was being done correctly.

The program has not been explicitly compared with actual data for any system to see if this simulation actually reproduces the reliability behavior of a water distribution system. It does have high face-validity since SDP8 is an accepted program for analyzing the hydraulic behavior of such a system. Also, as explained in Section 4.4, the failure and repair time distributions are "reasonable." However, the reliability results have never been compared with any real-world data, because a data base of information of this type is not available. Such validation would be desirable in the future.

4.6RESULTS

A printout of the input data for System A is presented in Table 4-6. Critical and minimum heads for this system were specified to correspond to 20 psi and 40 psi respectively. (Head limit (ft) = elevation (ft) + pressure (psi) * 2.307.) Various simulation runs were made for this system.No insurmountable problems were

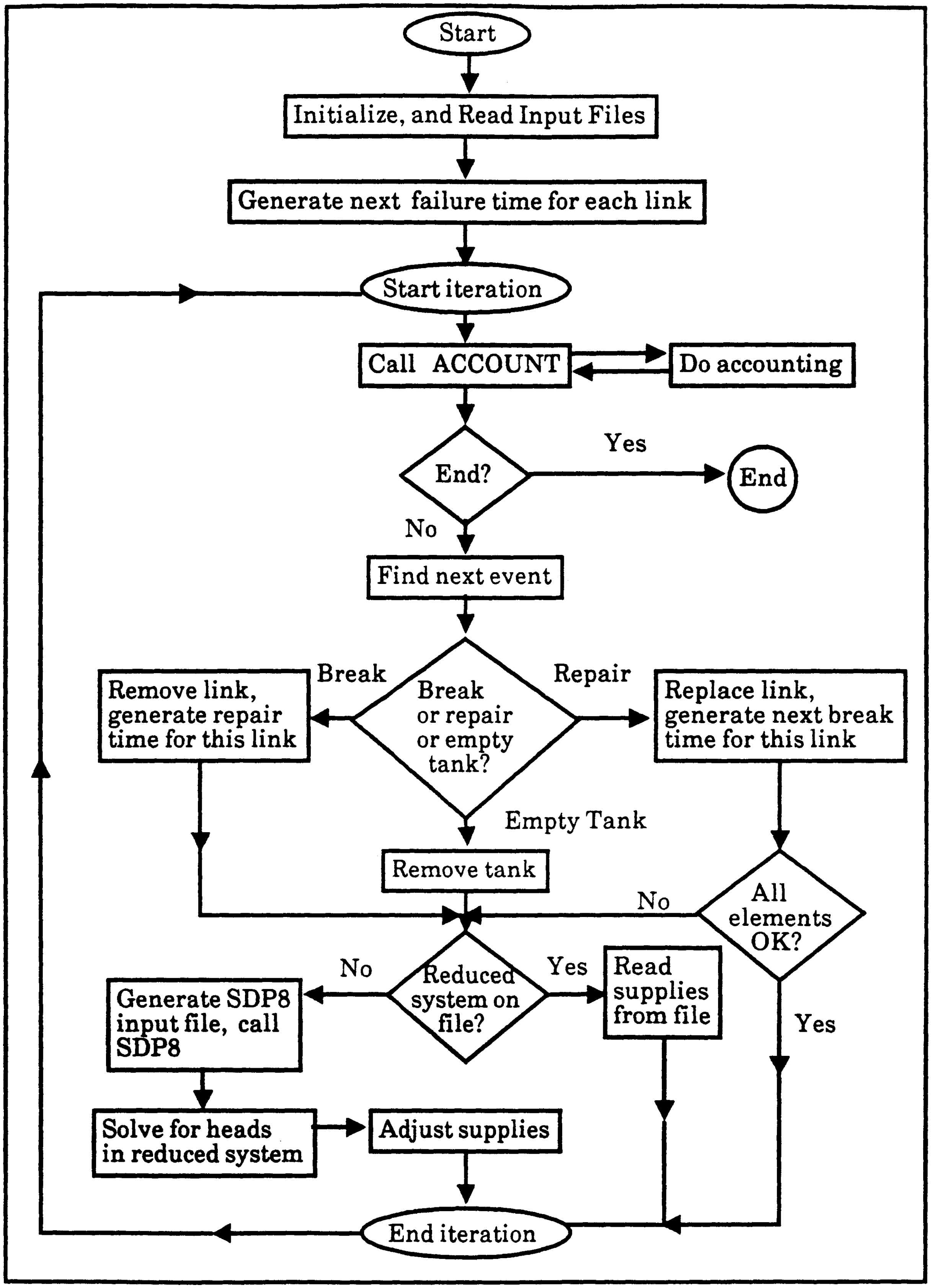


Figure 4-4 Simplified Flowchart

encountered in the simulation of System A. Highlights of the results of these simulations runs are presented below.

Table 4-7 shows the state of each node while each link is the only broken element in the system. During the entire two-hundred-year simulation run for System A, there was never a case where two pipes failed at the same time while the pump was operating. Thus all emergency conditions recorded in the simulation were caused by breaks in pipe 2 (connecting nodes 3 and 4) and pipe 5 (connecting nodes 5 and 6).

A summary of the results for System A is presented in Table 4-8. In this system there is very little problem with reduced service. Node 9 endures reduced service most often, but this mode occurs on average only for about one day per year and less than once per year. Failure conditions do occur relatively frequently. Note, all of the node failure results are very close to those of the pump. This correspondence indicates pump failures are the major source of unreliability in this system, as expected. It should be noted that the probability distribution parameters were consciously chosen to be high, so fairly high failure occurrences are not surprising.

Node Data:

Node	Elevation (feet)	Demand for Fully Working System (MGD)	Supply for Fully Working System (MGD)	Head for Fully Working System (feet)	Service Head Limit (feet)	Minimum Head Limit (feet)
1 (Res.)	100	-	6.625	100.00	-	-
2	100	1.6	-	388.48	192.28	146.14
3	200	1.2	-	386.43	292.28	246.14
4	210	0.6	-	376.80	302.28	256.14
5	230	0.4	-	377.54	322.28	276.14
6	250	0.825	-	380.05	342.28	296.14
7	10	0.6	-	173.57	102.28	56.14
8	10	0.8	-	170.31	102.28	56.14
9	50	0.4	-	160.87	142.28	96.14
10	25	0.2	-	181.37	117.28	71.14

Pump Data:

Pump curve: discharge = input head + 375 − $0.015^{*}\text{flow}^{4.58}$
Mean time to pump break: 1000 hr
Pump Repair Parameters for log-normal distribution: $\upsilon = 3.93$, $\sigma = 0.2$

Pipe Repair Data:

Uniform Distribution: 3 − 72 hours

Table 4-6 System A Input Data

Link Data:

Link	From	To	Length (feet)	Diameter (inches)	Hazen-Williams Coef.	Mean TIme to Pipe Break (hr)
1	2	3	200	16	120	231000
2	3	4	1500	12	120	30800
3	3	6	1800	14	120	25700
4	5	4	2000	10	120	23100
5	6	5	1900	14	120	24300
6	7	8	1000	8	120	46300
7	8	9	2500	10	120	18500
8	7	9	3500	8	120	13200
9	10	7	1500	10	120	30800
10	10	7	1500	6	120	30800
98 (valve)	4	10	500	6	65	92500
99 (valve)	5	8	500	4	65	92500
100 (pump)	1	2	-	-	-	-

Table 4-6 System A Input Data (continued)

Link	From	To	Nodes in Normal Status	Nodes in Reduced Service	Failed Nodes
1	2	3	1, 2		3, 4, 5, 6, 7, 8, 9, 10
2	3	4	1, 2, 3, 4, 5, 6, 7, 8, 10	9	
3	3	6	1, 2, 3, 4, 7, 8, 10	5, 6, 9	
4	5	4	all		
5	6	5	all		
6	7	8	all		
7	8	9	all		
8	7	9	all		
9	10	7	1, 2, 3, 4, 5, 6, 10	8, 7	9
10	10	7	all		
98	4	10	1, 2, 3, 4, 5, 6		7, 8, 9, 10
99	5	8	1, 2, 3, 4, 5, 6		7, 8, 9, 10
100	1	2			all

Table 4-7 Node Status for Single Link Failures

Normal: 40 psi $\leqq$ node pressure

Reduced service: 20 psi $\leqq$ node pressure $\leqq$ 40 psi

Failed: node pressure $\leqq$ 20 psi

System Results:

2350 cycles

4846 events

Average annual shortfall: 5.04%

121.08 million gallons (± 5.5)

Node Results:

Node	Average Time in Reduced Service			Average Time in Failure			Average Shortfall	
	%	hours / year	number /year	%	hours / year	number / year	%	annual average (10^6 gal)
1 (Res.)	0.0	0	0	0.0	0	0	-	-
2	0.0	0	0	4.98	436	8.4	4.97	29
3	0.0	0	0	4.99	437	8.4	4.99	22
4	0.0	0	0	4.99	437	8.4	4.99	11
5	0.15	13	0.4	4.99	437	8.4	4.99	7
6	0.15	13	0.4	4.99	437	8.4	5.03	15
7	0.14	13	0.3	5.07	444	8.6	5.10	11
8	0.14	13	0.3	5.07	444	8.6	5.10	15
9	0.28	25	0.7	5.21	457	8.9	5.27	8
10	0.0	0	0	5.07	444	8.6	5.07	4

Note – Confidence Intervals:

% in reduced service, within ± 0.05

% in failure, within ± 0.23

Pump Results:

Average Time in Failure: 4.97%

436 hours/year

8.4 / year

Table 4-8 System A Simulation Results

Link Results:

Pipe	Time in Failure		
	%	hours / year	number / year
1	0.02	1	0.04
2	0.15	13	0.3
3	0.16	14	0.4
4	0.14	12	0.3
5	0.14	12	0.3
6	0.11	10	0.2
7	0.25	22	0.6
8	0.33	29	0.7
9	0.14	13	0.3
10	0.12	10	0.3
98 (valve)	0.03	3	0.9
99 (valve)	0.05	4	0.1

Table 4-8 System A Simulation Results (continued)

On average then, the system appears to function completely about 95% of the time. The corresponding figures from the analytical methods in Chapter 3, are: connectivity = 95.4% and probability of sufficient flow = 94.26%. (These figures correspond to the "realistic" probability case [Case IV in Table 3-8].) The results of the methods agree closely for this system. However, it should be noted that in both cases these figures mainly reflect the reliability of the pump. For another comparison between the analytical and the simulation methods, we can compare the relative reliabilities of the individual nodes. From the reachability calculations in Table 3-2 node 1 is expected to be the most reliable and node 9 the least. The simulation results, both for reduced service and failure modes, agree. Note, although we might expect the nodes at the higher elevations to be the least reliable, for this system the least reliable nodes are those furthest from the source but at lower elevations.

Since the pumps are not most important determinant of reliability in this system, a number of changes in the pump related parameters were examined as options for system improvement. The alternatives simulated were:

(1) improvement of pump maintenance, resulting in an increase of mean time between failures from 1000 to 1500 hours,
(1) improvement of pump repair, resulting in an decrease of mean time to repair from 50 to 40 hours,
(3) a combination of alternatives 1 and 2.

Figure 4-5 presents, for the base case (B) and the improvement options (1, 2, and 3), the calculated values of annual shortfall and percentage of time in each mode for node 9. Node 9 was chosen for these comparisons, because it is the node with the most severe reliability problems. Note, increasing the maintenance to the pump causes a larger decrease in shortfall and failure percentage at node 9, than does improving the pump repair time. The combination of these two options decreases shortfall and percentage of time in failure mode by approximately 45%.

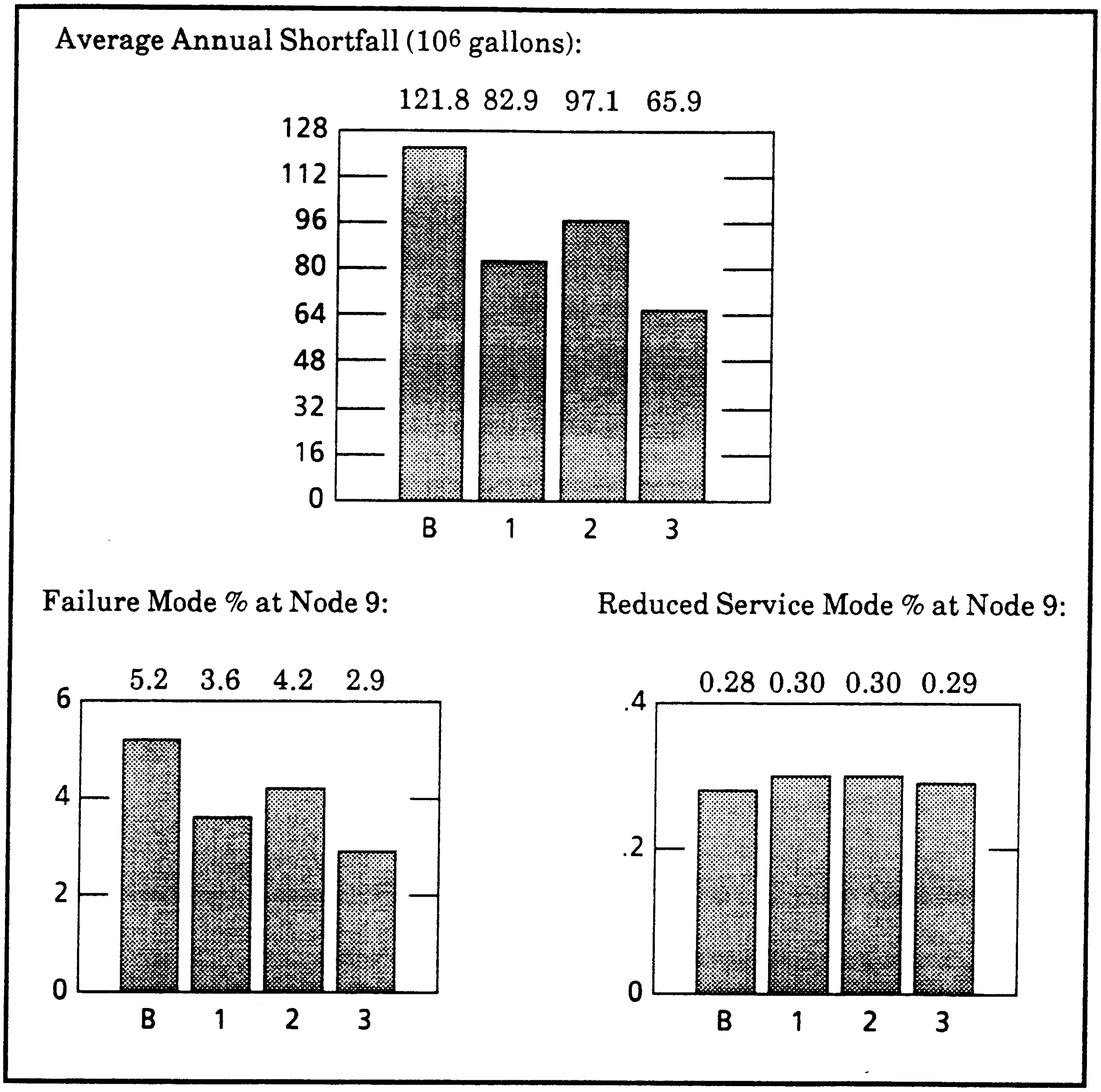


Figure 4-5 Comparison of Results for System A Options

Another improvement option could be to add another pump to the system. A smaller pump connected in parallel could be an economical way of both increasing the normal effectiveness of the system, and providing added reliability. However, using SDP8 alone shows adding a smaller pump in parallel is unlikely to augment the system reliability. A pump with 1/3 the capacity was tried, but this pump could not supply the network on its own. Thus most nodes will still be out the 5% of the time the large pump is inoperative. A pump of 1/2 the size of the original pump was also tried, but this pump alone could not supply the network either. Obviously the reliability of this system could be enhanced by adding another full sized pump, either in parallel to the first or as a standby. However, this option is likely to be very costly.

A listing of the input for System B+ is given in Table 4-9. The critical and minimum heads were set to be equivalent to pressures of 40 psi and 10 psi, respectively. The 40 psi is design standard for this system; and 10 psi is considered to be the pressure at which it is reasonable to cut off water altogether.

This system is quite large, and frequently more than one link failed at the same time. Also the tanks are frequently depleted. Thus it is hard to tell which configurations to pre-compile. Simulating all the 2 link failure combination involves $37^2 = 1369$ SDP8 solutions, which requires unreasonable computer time (several CPU hours). However, the simulation runs themselves take considerable computer time, so some savings due to pre-compilation of data is desirable. As a compromise, runs were made and data stored of:

(1) each pipe failed individually (34 runs),
(2) one pump failed (1 run),
(3) one pump failed <u>and</u> each pipe failed individually (34 runs),
(4) tank 165 failed <u>and</u> each each pipe failed individually (34 runs),
(5) both tanks failed <u>and</u> each pipe failed individually (34 runs).

For both computer time and storage requirements a simulation period of 3 years (approximately 140 regeneration cycles) was chosen. This period gave confidence

Node Data:

Node	Elevation (feet)	Demand for Fully Working System (gpm)	Supply for Fully Working System (gpm)	Head for Fully Working System (feet)	Service Head Limit (feet)	Minimum Head Limit (feet)
10 (river)	10	-	4428	10	-	-
20	20	500	-	305.61	112.28	43.07
30	50	200	-	242.40	142.28	73.07
40	50	200	-	234.41	142.28	73.07
50	50	200	-	232.16	142.28	73.07
60	50	500	-	234.96	142.28	73.07
65 (tank)	235	-	342	235	-	-
70	50	500	-	242.40	142.28	73.07
80	50	500	-	228.33	142.28	73.07
90	50	1000	-	225.88	142.28	73.07
100	50	500	-	229.33	142.28	73.07
110	50	500	-	234.50	142.28	73.07
120	120	200	-	228.67	212.28	143.07
130	120	200	-	228.69	212.28	143.07
140	80	200	-	228.34	172.28	103.07
150	120	200	-	228.32	212.28	143.07
160	120	800	-	234.27	212.28	143.07
165 (tank)	235	-	1630	235	-	-
170	120	200	-	226.10	212.28	143.07

Pump Data (same for all 3 pumps):

Pump curves: discharge = input head + 300 − $2.43 \times 10^{-6} * \text{flow}^{1.974}$

Mean time to pump break: 1000 hr

Pump Repair Parameters for log-normal distribution: $\upsilon = 3.93, \sigma = 0.2$

Table 4-9 System B+ Input Data

Pipe Repair Data:

Uniform Distribution: 3 – 72 hours

Tank Data:

Diameter: 100 ft
Initial Depth: 10 ft
InitialVolume: 78,540 gallons

Link Data:

Link	From	To	Length (feet)	Diameter (inches)	Hazen-Williams Coef.	Mean Time to Pipe Break (hr)
2	20	70	12000	16	70	3855
4	20	30	12000	12	120	3855
6	20	110	12000	12	70	3855
8	30	70	9000	12	70	5140
10	70	100	6000	12	70	7710
12	70	90	6000	10	70	7710
14	70	60	6000	12	70	7710
16	90	60	6000	10	70	7710
18	60	80	6000	12	70	7710
20	90	80	6000	10	70	7710
22	90	150	6000	10	70	7710
24	90	100	6000	10	70	7710
26	100	150	6000	12	70	7710
28	80	150	6000	10	70	7710
30	30	60	6000	10	120	7710

Table 4-9 System B + Input Data (continued)

Link Data (continued):

Link	From	To	Length (feet)	Diameter (inches)	Hazen-Williams Coef.	Mean Time to Pipe Break (hr)
32	30	40	6000	10	120	7710
34	30	50	6000	10	120	7710
36	40	50	6000	10	120	7710
38	80	50	6000	10	120	7710
40	80	140	6000	10	120	7710
42	150	140	6000	8	120	7710
44	150	160	6000	8	120	7710
46	100	160	6000	8	120	7710
48	100	110	6000	8	70	7710
50	110	160	6000	10	120	7710
52	110	120	6000	8	120	7710
56	120	130	6000	8	120	7710
58	160	130	6000	10	120	7710
60	130	170	6000	8	120	7710
62	160	140	6000	8	120	7710
64	170	140	12000	8	120	3855
66	140	50	12000	8	120	3855
70	60	65	100	12	120	46260
80	160	16520	100	12	120	46260
101 (pump)	10	20	–	–	–	–
102 (pump)	10	20	–	–	–	–
103 (pump)	10	20	–	–	–	–

Table 4-9 System B+ Input Data (continued)

intervals for System B+ to within 18% for shortfalls. Again the confidence intervals for the small percentages of time each node was in failure or reduced service mode were sometimes quite wide.

Various simulation runs were made for this system. Highlights of the results of these simulations runs are presented in Table 4-10.

The nodes with the high percentages of time in reduced service and failure mode, are the nodes that cannot be supplied by the pumps alone. These percentages correspond primarily to the 20% of the time both tanks have been depleted. Walski reports the system has trouble filling tank 165, so it does not seem surprising that 20% of the time (about 1.5 days/week), there is some supply problem at the nodes at the higher elevations (nodes 120, 130, 160, and 170). Walski did not specify the diameter of the tanks; these diameters were picked as "reasonable".

Other studies have been made on this system (Lee *et. al.* [1985], Gessler [1985]), however they focused on finding a design to meet projected future water demand rather than on reliability. Lee used 800,000 gallon tanks which were not allowed to run dry. Gessler used a tank with 200,000 of available flow. These tanks are approximately in the range of the 587,000 of available water used in this simulation. However in both of these studies, the recommended expansion involved another tank added to the system.

It seems reasonable that in designing storage for this system, the volume of the water storage tanks should take into account the amount of water needed to supply the system during failure events. The required tank volume should be related to the distribution of the pipe repair time. Thus, for improvement

System Results:

2350 cycles
4846 events

Average annual shortfall:	7.39%	
	248.68 million gallons (± 44.7)	

Pump Results:

Pump 101:		
	Average Time in Failure:	4.51%
		395 hours/year
		7.3 / year
Pump 102:		
	Average Time in Failure:	4.03%
		353 hours/year
		7.0 / year
Pump 103:		
	Average Time in Failure:	3.93%
		344 hours/year
		7.0 / year

Table 4-10 System B+ Simulation Results

Node Results:

Node	Average Time in Reduced Service			Average Time in Failure			Average Shortfall	
	%	hours / year	number / year	%	hours / year	number / year	%	annual average (10^6 gal)
10 (Riv)	0.0	0	0	0.0	0	0	-	-
20	0.0	0	0	0.0	0	0	0.00	0
30	0.66	57	2.7	1.39	120	3.3	1.59	1.7
40	1.33	115	3.0	1.62	140	4.7	1.89	2.0
50	11.16	965	23.0	1.62	140	4.7	2.20	2.3
60	9.96	862	20.3	1.62	140	4.7	2.10	5.5
65 (Tk)	0.0	0	0	20.02	1734	43.0	-	-
70	0.66	57	2.7	1.39	120	3.3	1.60	4.2
80	18.34	1588	41.0	1.62	140	4.7	3.77	9.9
90	18.34	1588	41.0	1.62	140	4.7	4.12	21.7
100	18.34	1588	41.0	1.62	140	4.7	3.92	10.3
110	18.08	1566	41.0	1.62	140	4.7	3.24	8.5
120	2.13	186	42.3	20.01	1733	43.3	20.19	21.2
130	2.18	190	42.3	20.01	1733	43.3	20.20	21.2
140	17.78	1540	43.3	2.23	193	6.7	9.81	10.3
150	2.17	190	43.7	19.96	1729	43.0	19.99	21.0
160	2.17	190	43.7	19.96	1729	43.0	20.15	84.7
165 (Tk)	0.0	0	0	22.14	1919	44.7	-	-
170	2.22	194	43.7	19.96	1729	43.0	20.21	21.2

Note — Confidence Intervals:

% in reduced service, within ± 3.1

% in failure, within ± 3.4

Table 4-10 System B+ Simulation Results (continued)

Link Results:

Pipe	Time in Failure		
	%	hours / year	number / year
2	0.56	49	1.0
4	1.02	89	2.3
6	0.34	30	1.7
8	0.25	22	0.7
10	0.37	33	0.7
12	0.37	33	0.7
14	0.55	48	1.0
16	0.50	44	1.7
18	0.41	36	0.7
20	0.45	39	1.7
22	0.31	27	1.0
24	0.50	44	1.0
26	0.50	43	1.0
28	0.33	29	1.7
30	0.35	31	0.7
32	0.72	63	1.0
34	0.69	60	1.7
36	0.63	56	1.3
38	0.77	67	1.3
40	0.05	5	0.3

Table 4-10 System B+ Simulation Results (continued)

Link Results (continued):

Link	Time in Failure		
	%	hours / year	number / year
42	0.63	56	1.3
44	0.74	65	1.3
46	0.73	64	2.0
48	0.43	38	1.3
50	0.38	33	0.7
52	0.19	17	0.7
56	0.11	10	0.3
58	0.32	28	0.7
60	0.33	29	1.0
62	0.75	66	1.7
64	0.89	78	2.0
66	1.14	100	3.0
70	0.47	41	1.0
80	0.11	9	0.7

Table 4-10 System B + Simulation Results (continued)

alternatives, combinations of pipe repair time distributions and tanks sizes were examined. The following options were examined:

(1) tank diameter = 100 ft, pipe repair time = 3 - 72 hours (base case),
(2) tank diameter = 150 ft, pipe repair time = 3 - 72 hours,
(3) tank diameter = 100 ft, pipe repair time = 3 - 48 hours,
(4) tank diameter = 150 ft, pipe repair time = 3 - 48 hours,
(5) tank diameter = 100 ft, pipe repair time = 3 - 24 hours, and
(6) tank diameter = 150 ft, pipe repair time = 3 - 24 hours.

A summary of the shortfalls for each combination is given in Figure 4-6. The "best" alternative (large tanks and pipe repairs within 24 hours) gives an average annual shortfall of 93 ± 22 million gallons. For this case the maximum time in reduced service mode is 7.3 ± 1.7 % for node 110, and the maximum time in failure mode is 7.6 ± 1.7 % for nodes 120 and 130.

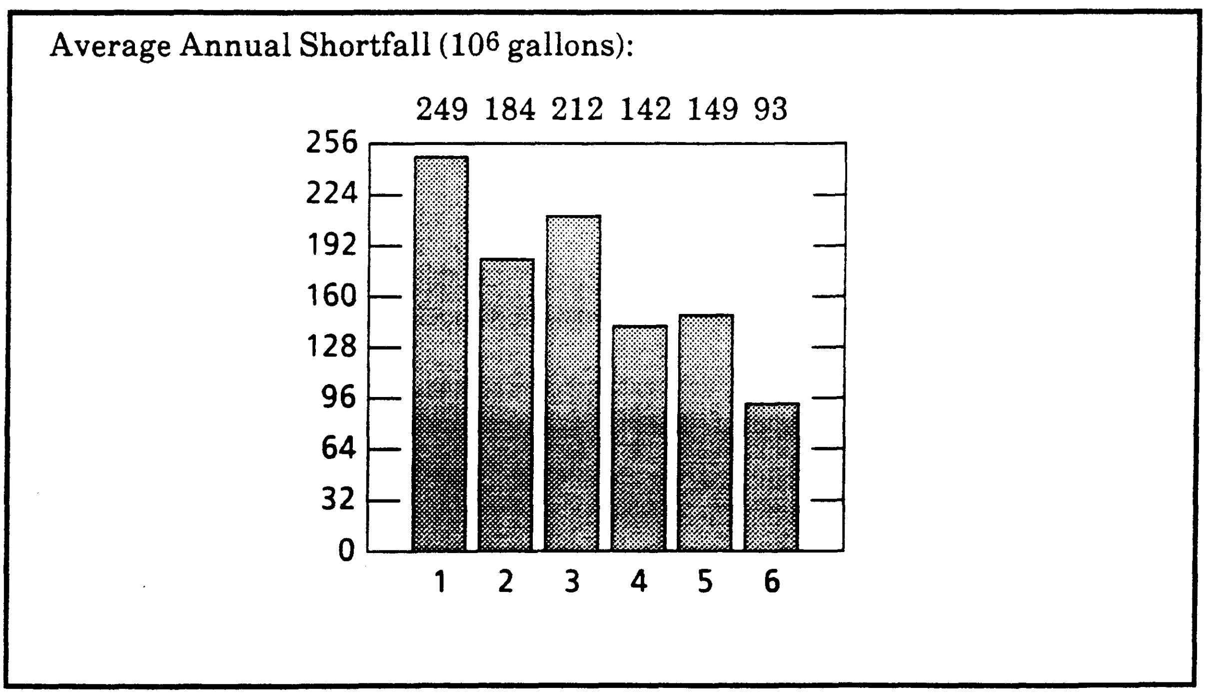


Figure 4-6 Comparison of Results for System B+ Options

To allow a comparison with the results of the analytical computations in the previous chapter, one simulation run was made on Network B (Figure 3-10). A listing of the input for Network B is given in Table 4-11. Analogously to the large network, System B+, the critical and minimum heads were set to be equivalent to 40 psi and 10 psi, respectively. Since only one simulation run was planned, no preproccessing was performed for this system. Thus, every time the simulation encountered a system with any element failed SDP8 was called.

The simulations for this system, with no prestored supply results, required 368 CPU seconds. Highlights of the results of this simulation are given in Table 4-12.

Within this three year simulation period, 79 cycles occurred, encompassing 244 events. This period gave confidence intervals to within 23% for shortfalls. Again, this system has severe problems with reliability. The tank failed (was dry) for 11.7% of the simulation period. Node 150 was the least reliable node, failing the same 11.7% of the time the tank was dry. The nodes in the pentagon failed approximately 2% of the time. The nodes in the pentagon also spent significant time in reduced service mode, about 10%. Again the nodes furthest away from the water sources have the most reliability problems, not a surprising result since all of these nodes (except 20 and 150) are at the same elevation.

From Chapter 3, the system connectivity is 0.9534 and the probability of sufficient supply is 0.9514. These figures, however, assume that the tank is completely reliable and thus never runs dry. A comparison with the simulation results show that these analytical measures are somewhat optimistic for this system, as expected.

Node Data:

Node	Elevation (feet)	Demand for Fully Working System (gpm)	Supply for Fully Working System (gpm)	Head for Fully Working System (feet)	Service Head Limit (feet)	Minimum Head Limit (feet)
10 (river)	10	-	3026	10	-	-
20	20	500	-	306.09	112.28	43.07
60	50	500	-	234.60	142.28	73.07
65 (tank)	235	-	1174	235	-	-
70	50	500	-	238.62	142.28	73.07
80	50	500	-	217.43	142.28	73.07
90	50	1000	-	216.97	142.28	73.07
100	50	500	-	220.12	142.28	73.07
110	50	500	-	248.57	142.28	73.07
150	120	200	-	217.30	212.28	143.07

Pump Data (same for both pumps):

Pump curves: discharge = input head + 300 − $2.43 \times 10^{-6} * \text{flow}^{1.974}$

Mean time to pump break: 1000 hr

Pump Repair Parameters for log-normal distribution: $\upsilon = 3.93, \sigma = 0.2$

Pipe Repair Data:

Uniform Distribution: 3 − 72 hours

Tank Data:

Diameter: 100 ft

Initial Depth: 10 ft

InitialVolume: 78,540 gallons

Table 4-11 Network B Input Data

Link Data:

Link	From	To	Length (feet)	Diameter (inches)	Hazen-Williams Coef.	Mean Time to Pipe Break (hr)
2	20	70	12000	16	70	3855
6	20	110	12000	12	70	3855
10	70	100	6000	12	70	7710
12	70	90	6000	10	70	7710
14	70	60	6000	12	70	7710
16	90	60	6000	10	70	7710
18	60	80	6000	12	70	7710
20	90	80	6000	10	70	7710
22	90	150	6000	10	70	7710
24	90	100	6000	10	70	7710
26	100	150	6000	12	70	7710
28	80	150	6000	10	70	7710
48	100	110	6000	8	70	7710
78	60	65	100	12	120	46260
101 (pump)	10	20	-	-	-	-
102 (pump)	10	20	-	-	-	-

Table 4-11 Network B Input Data (continued)

System Results:

79 cycles

244 events

Average annual shortfall: 4.62%

101.98 million gallons ($\pm$ 23.4)

Node Results:

Node	Average Time in Reduced Service			Average Time in Failure			Average Shortfall	
	%	hours / year	number / year	%	hours / year	number / year	%	annual average (10^6 gal)
10 (Riv)	0.0	0	0	0.0	0	0	–	–
20	0.0	0	0	0.0	0	0	0.00	0
60	9.63	843	20.7	2.09	183	5.3	6.23	16
65 (Tk)	0.0	0	0	11.72	1027	25.3	–	–
70	7.01	614	14.7	1.01	88	3.0	1.69	4
80	9.63	843	20.7	2.09	183	5.3	6.93	18
90	9.63	843	20.7	2.09	183	5.3	6.71	35
100	10.42	912	22.3	1.30	114	4.0	4.78	13
110	0.09	8	1.7	1.01	88	3.0	1.04	3
150	0.45	39	6.3	11.72	1027	25.3	11.76	12

Note – Confidence Intervals:

% in reduced service, within $\pm$ 3.1

% in failure, within $\pm$ 3.4

Table 4-12 Network B Simulation Results

Pump Results:

Pump 101:

Average Time in Failure: 4.87%
427 hours/year
8.7 / year

Pump 102:

Average Time in Failure: 3.23%
283 hours/year
5.7 / year

Link Results:

Pipe	Time in Failure		
	%	hours / year	number / year
2	0.59	52	2.7
6	0.58	51	2.7
10	0.35	31	0.7
12	0.20	18	0.3
14	1.06	93	2.0
16	0.69	61	1.3
18	0.14	13	0.7
20	0.57	50	1.3
22	0.50	44	1.0
24	0.39	34	0.7
26	0.53	47	1.3
28	0.38	33	0.7
48	0.20	17	0.3
78	0.22	20	0.3

Table 4-12 Network B Simulation Results (continued)

4.7 DISCUSSION

As shown by the previous examples, simulation can be a useful tool for reliability assessment. Although simulation seems a time consuming task in comparison with the analytical methods presented earlier, in this context simulation provides three advantages.

First, with simulation a number of reliability measures can be calculated. As shown in Table 4-2 this program already calculates a number of measures. With only minor modification the program could record additional measures such as, the duration of the longest period of failure at any node, the duration of the longest period of reduced service at any node, and the period in which the greatest total shortfall occurred. Only with simulation is such flexibility in reliability criteria possible.

Second, simulation allows the analysis of a system with complicated interactions. This system included operational response to supply loss, water tanks with storage dependent on the state of the system, and fairly detailed modeling of the reliability of the individual pipes in the system. Analytical methods have been designed which handle, to some extent, some of such complexities. However, to analyze a system with all of these elements at once requires simulation. Simulation provides a level of realism available with no other method.

Third, simulation allows the detailed modeling of the hydraulic behavior of the system. In contrast, to remain tractable most analytical methods require a simplified description of the water system. In chapter 3, considerable effort was spent determining how well the model of capacitated flow could be "forced" on the water system. By using simulation with an accepted model of hydraulic behavior in a piping network, like SDP8, the risk of a totally incorrect response in some untested range of the model can be lessened.

However as previously mentioned, simulation can be time consuming, both in terms of computer time required per analysis and in terms of time to set up and use such a program. Also simulation results are hard to optimize, and can be hard to generalize beyond a very specific system. Thus perhaps the best approach to a reliability analysis is to use both simulation and analytical of methods. Exactly how these two approaches can be used together is discussed next.

5.0 CONCLUSION

5.1 OVERVIEW OF RELIABILITY ASSESSMENT OF WATER DISTRIBUTION SYSTEMS

Several methods for calculating reliability measures for water distribution systems have been presented in previous chapters. For each method, calculations have been performed for sample systems. In this section, an overview of these methods is presented, in order to suggest how they may be used together towards the development of a comprehensive understanding of the reliability of an existing system. This methodology does, of course, rely on the previously presented methods, however it also relies on an equally important component of any engineering analysis – sound engineering judgment.

As a means of building understanding and developing an intuitive sense of the hydraulic behavior of the system, we recommend as a first step in any reliability analysis the investigation of several network representations with the a computer model of the hydraulic behavior of the system. The simulation program presented in Chapter 4 used such a model, SDP8, as a subroutine. However, before the simulation is run, and frequently before any analytical calculations are done, it may be useful to use the hydraulic model as a stand-alone program. At a minimum, the following system configurations should be investigated in an initial analysis:

* all components operational,
* each water source in turn inoperative,
* each pump in turn inoperative, and
* frequently expected combinations of pumps and sources inoperative.

For a system with which the analyst is already familiar, the use of a computer model to understand the system may be unnecessary. However, initial use of the network solver with the proposed network representation can also be used to investigate how well this system model actually describes the real system. Execution of this initial step can also lead to the identification and elimination of glaring errors in such a representation.

As the next step of the reliability assessment, we recommend the use of relevant analytical methods. As was shown in Chapter 3, measures of connectivity and reachability are fairly easy to calculate even for moderately sized, complex systems. These measures can be used to identify basic sources of unreliability in a system such as lack of network interconnections or extremely unreliable links. In addition, nodes with reachabilities below those of others in the system may be initially identified as problem areas in the system. Following these relatively easy calculations, it is also important to calculate the more complex measures of probabilistic supply. As described in Chapter 3, some difficulties were encountered with the calculation of these measures even on moderately sized, complex networks. However, either by using a simplified representation of the network in question, or by obtaining faster and more sophisticated algorithms and software for these calculations, these probabilistic supply calculations should be possible for most water distribution systems. These measures will identify nodes in the network which, although connected, do not reliably receive the amount of water demanded.

Armed with the above information and calculations, the analyst can perform a tentative reliability assessment. If these analyses, or past experience, indicate an existing reliability problem these calculations should help to pinpoint specific areas of concern. Once problem areas are identified, possible improvements to the system may be suggested.

To some extent, the improvement options can be investigated with the methods mentioned thus far. For example, if a new link is suggested, it would be worthwhile to re-run some of the system configurations using the hydraulic model, and to recalculate the connectivity, reachabilities, and probabilistic supply measures for the augmented system. We envision the above steps being done in an iterative fashion, until promising improvements are identified.

When a small number of improvement options have been suggested, simulation (as in Chapter 4) can be used to investigate these alternatives more fully. In the simulation analysis, as many of the complexities of the system as possible should be included. The simulation in Chapter 4 included elements such as water tanks, and operational response to shortages in supply. For other systems additional features may be desirable. Also the simulations should be designed so as to calculate a wide range of reliability measures, including the measures previously calculated analytically, as a check on the accuracy and agreement of the various approaches. If done carefully and with good judgment, simulations can provide a very good understanding of the reliability of these alternative systems under real-life conditions. Again, the lest two steps may proceed iteratively, with simulation results suggesting new alternatives and vice versa.

In conclusion, previous chapters present a number of reliability assessment tools. These tools, combined with care, experience, and common sense can be used to provide a detailed and fairly comprehensive assessment of the reliability of an existing system, and can be used to compare alternative schemes for improving that system. An outline of the steps proposed in this subsection is presented in Figure 5-1.

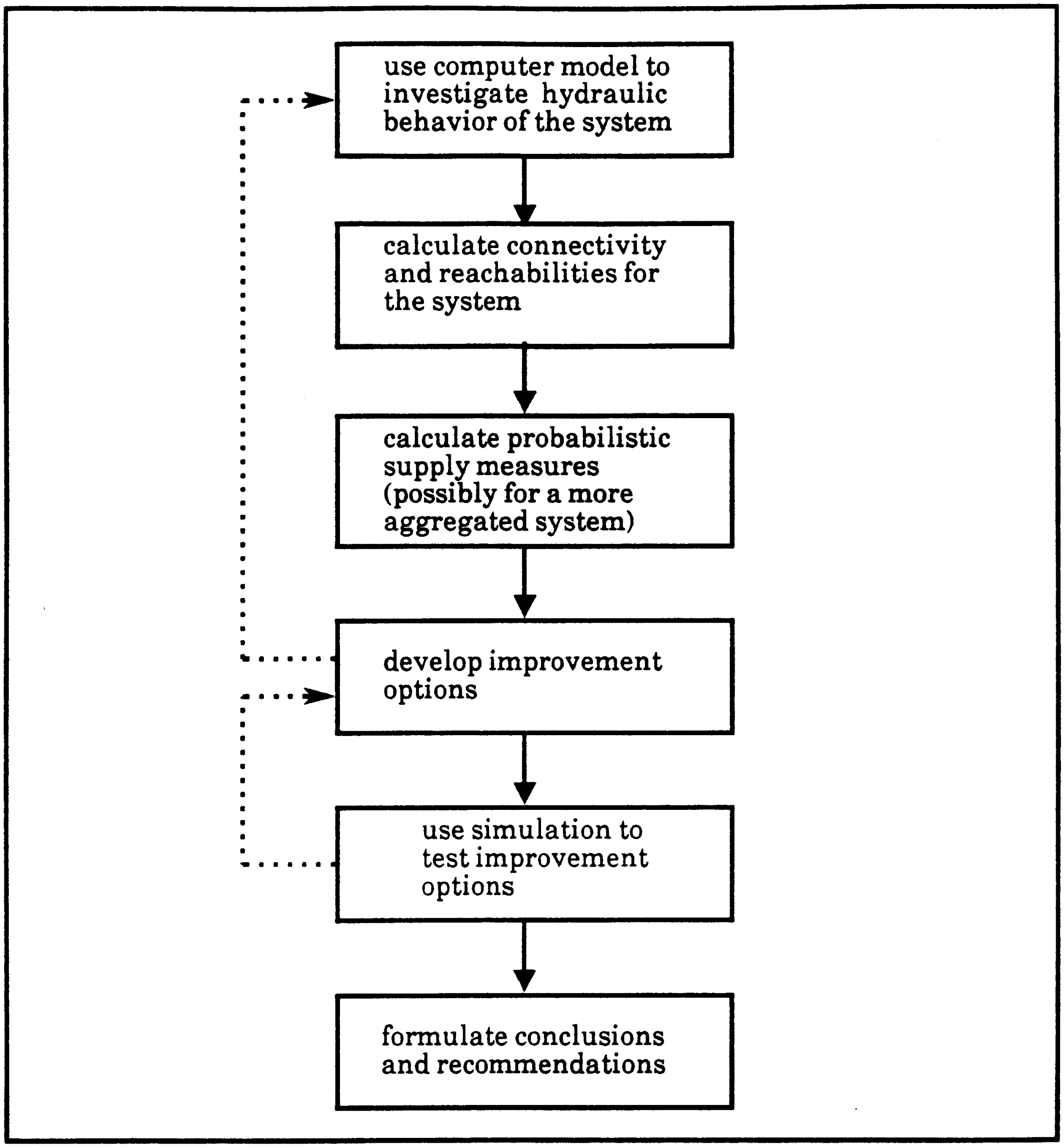


Figure 5-1 Proposed Process for Reliability Assessment

5.2 SUGGESTIONS FOR FURTHER WORK

There are a number of projects and areas for further research suggested by this work. This additional effort can be classified into two areas:

* theoretical and algorithmic development, and
* software and computer system development.

Of the methods discussed in this report, those for calculating probabilistic flow measures most need additional theoretical development. As was discussed in Chapter 3, the program developed using Lee's algorithm [1980] required large amounts of computer time for even moderately sized systems. However, as mentioned in Chapter 2, other more sophisticated algorithms for these measures exist. Willie [1979] presents what appears to be a faster method for calculating reliability of capacitated flow networks. Part II of his paper discusses a flow network analysis program he developed for these algorithms. If this code is publicly available, it would be interesting to obtain it and apply it to water distribution networks. Even if this code is not available, the development of another program for the calculation of probabilistic flow measures, using the outline of Willie's algorithm as presented in his paper, would be of value. Alternatively, Rosenthal's algorithm, employed in Chapter 3, has also been extended to the calculation of reliability measures on capacitated networks. Since this algorithm seemed to work well for planar networks it also is likely to reduce the computations required for these measures. Additional investigations of the adequacy of the methods used in Chapter 3 for assigning link capacity values should also be performed.

Once the tools to calculate the measures investigated in previous chapters are refined, a number of further analyses could be performed. It would be interesting to study the relationship between these reliability measures and other characteristics of the network. Such characteristics include the topology of the network, the values of the reliabilities of the individual components, and the location and amount of water storage in the network. Additionally, Lee's algorithm can be used with other definitions of an operational system. For example a system could be said to be operational if only some of the nodes receive sufficient supply, instead of all the nodes as was previously used. Changing the definition of an operating system then changes the reliability measures calculated based on this definition. It would be interesting to see how such changes affect the reliability values calculated for a given system.

Another useful theoretical development suggested by this work, would be to use stochastic gradient techniques in addition to the simulation developed in Chapter 4. Such an extension would then lead to formal ways of finding improvement options and can be used to optimize the system with respect to reliability considerations. It should also be possible to relate the improvements to cost. A tradeoff curve of reliability versus cost could be developed by using stochastic gradient techniques to optimize the system for a number of proposed funding levels. Joining cost, reliability, and other design considerations can lead to true multi-objective design methods for water distribution networks.

The largest need, however, if the ideas and methods in this report are to be used by practitioners, is to develop a more integrated system of computer software. For the analyses in previous chapters, an *ad hoc* conglomeration of programs was developed with little attempt to make these programs user-friendly, or even usable by someone not familiar with them. For example, although the programs use similar information about the system configuration and other parameters relevant to the reliability of the piping system, most of these programs now require separate data files. Additionally, software for calculations of reachability and connectivity is either publicly available, or could easily be developed, to perform the calculations done by hand in Chapter 3.

One can envision a system of integrated software for the assessment of water distribution system reliability, based on the methods presented in this report. Such a system would have easy editing and updating capabilities for managing the required inputs to, and for storing and integrating the outputs from, these programs. The programs themselves should be easy to use, and should run off common data files. In the best of all worlds, these programs would also have extensive display capabilities for showing the outputs in useful forms, including dynamic color graphics displays. It is even possible such a system could be developed to run on a personal computer, if the user was willing to allow the simulations to run for relatively long periods of time.

Even without all the capabilities described above, there is much that could be done with the software tools developed for this report, with perhaps some additional purchased software for reachability and connectivity calculations. Starting with these tools, a system for performing water system reliability assessments as outlined in Figure 5-1 could be developed that would be directly useful for planners and operators of such systems.

REFERENCES

Agrawal, A. and Barlow, R. E. 1984. An Survey of Network Reliability and Domination Theory. *Operations Research* 32 (1984):478-492.

Agrawal, A. and Satyanarayana, A. 1984. An O(|E|) Time Algorithm for Computing the Reliability of a Class of Directed Networks. *Operations Research* 32 (1984):493-515.

Aneja, Y. P., and Nair, K. P. K. 1980. Maximal Expected Flow in a Network Subject to Arc Failures. *Networks* 10 (1980): 45-57.

Ang, A. H-S. and Tang, W. H. *Probability Concepts in Engineering Planning and Design*. Volume II: Decision, Risk, and Reliability. New York: Wiley.

Arad, N. 1968. *A method for the evaluation of the system and cost-effectiveness of large multi-stage flash desalting plants*. Rep. R-1142. Washington D. C.: Plann. Res. Corp.

Ball, M. O. 1980. Complexity of Network Reliability Calculations. *Networks* 10 (1980): 153-165.

Barlow, R. E. 1984. Mathematical Theory of Reliability: A Historical Perspective. *IEEE Transactions on Reliability* R-33 (1984): 16-20.

Barlow, R. E., and Proschan, F. 1975. *Statistical Theory of Reliability and Life Testing: Probability Models*. New York: Holt, Rinehart, &Winston.

Billinton, R. 1972. Bibliography on Probability Methods in Power System Reliability Evaluation. *IEEE Transactions on Pwer Apparatus and Systems* PAS-91 (1972): 649-660.

Billinton, R., and Allan, R. N. 1983. *Reliability Evaluation of Engineering Systems: Concepts and Techniques*. New York: Plenum Press.

Billinton, R., and Allan, R. N. 1984. *Reliability Evaluation of Power Systems*. New York: Plenum Press.

Bradley, S. P., Hax, A. C., and Magnanti, T. L. 1977. *Applied Mathematical Programming*. Reading, MA: Addison-Wesley.

Brill, E. D., Liebman, J. C., and Lee, H-L. 1985. Optimization of Looped Water Distribution Networks . In *Computer Applications in Water Resources*, ed. H. Tourno, pp. 569-571. New York: ASCE.

Buzacott, J. A. 1980. A Recursive Algorithm for Finding Reliability Measures Related to the Connection of Nodes in a Graph. *Networks* 10 (1980): 311-327.

Carey, M. and Hendrickson, C. 1984. Bounds on Expected Performance of Networks with Links Subject to Failure. *Networks* 14 (1984): 439-456.

Damelin, E., Shamir, U., and Arad, N. Engineering and Economic Evaluation of the Reliability of Water Supply. *Water Resources Research* 8 (1972): 861-877.

De Jong, J. J., Snijders, J. A. C., and Wedmijer, A. 1983. Reliability of the Rotterdam drinking water system. (In Dutch) *H_2O* 16 (1983). Quoted in Tangena and Koster [1983].

Endrenyi, J. 1978. *Reliability Modeling in Electric Power Systems*. New York: Wiley, Interscience Publication.

Gessler, J. 1985. Pipe Network Optimization by Enumeration. In *Computer Applications in Water Resources*, ed. H. Tourno, pp. 572-581. New York: ASCE.

Henley, E. J., and Kumamoto, H. 1981. *Reliability Engineering and Risk Assessment*. Englewood Cliffs, NJ: Prentice Hall.

Hobbs, B. F. 1985a. Reliability Analysis of Water System Capacity. In *Hydrology and Hydraulics in the Small Computer Age*, ed. W. Waldrop, New York: ASCE, forthcoming.

----------. 1985b. Reliability Analysis of Urban Water Supply. In *Computer Applications in Water Resources*, ed. H. Tourno, pp. 1129-1238. New York: ASCE.

IEEE Subcommittee on the Application of Probability Methods, 1978. Bibliography on Probability Methods in Power System Reliability Evaluation. *IEEE Transactions on Pwer Apparatus and Systems* PAS-97 (1978): 2235-2242.

Johnson, R. 1984. Network Reliability and Acyclic Orientation. *Networks* 14 (1984): 489-505.

Law, A. M. and Kelton, W. D. 1982. *Simulation Modeling and Analysis*. New York: McGraw-Hill.

Lee, S. H. 1980. Reliability Evaluation of a Flow Network. *IEEE Transactions on Reliability* R-29 (1980): 24-26.

Lie, C. H., Hwang, C. L., and Tillman, F. A. Availability of Maintained Systems: A State-of-the-Art Survey. *AIEE Transactions* 9 (1977): 247-259.

Marks, D. H., Jeffrey, L., Andreou, S., and Park, C. 1985. *Predicting Urban Water Distribution Maintenance Strategies: A Case Study of New Haven, Connecticut*. Report to the Environmental Protection Agency by the Department of Civil Engineering, MIT, Cambridge, MA. January 1985.

Murchland, J. D. *Fundamental Concepts and Relations for Reliability Analysis of Multi-state Systems*. Edited by Barlow, R. E., Fussell, J. B., and Singpurwalla, N. D., pp. 581-618. Philedelphia: Society for Industrial and Applied Mathematics.

O'Day, K. Organizing and analyzing leak and break data for making main replacement decisions. *Journal AWWA* (1982): 588-594.

Provan, J. S., and Ball, M. O. 1983a. The Complexity of Counting Cuts and Computing the Probability that a Graph is Connected. *SIAM J. Comput.* 12 (1983): 777-788.

----------. 1983b. Calculating Bounds on Reachability and Connectedness in Stochastic Networks. *Networks* 13 (1983): 253-278.

----------. 1984. Computing Network Reliability in Time Polynomial in the Number of Cuts. *Operations Research* 32 (1984): 516-526.

Rosenthal, A. 1977. Computing the Reliability of Complex Networks. *SIAM J. Appl. Math.* 32 (1977): 384-393.

Satyanarayana, A., and Wood, R. K. 1982. Polygon-to-Chain Reductions and Network Reliability, ORC 82-4, Operations Research Center, University of California, Berkeley, CA

Shamir, U. and Howard, C. D. D. 1979. An Analytic Approach to Scheduling Pipe Replacement. *Journal AWWA* (1979): 248-258.

----------. 1981. Water supply reliability theory. *Journal AWWA* (1981): 379-384.

----------. 1985. Reliability and Risk Assessment for Water Supply Systems. In *Computer Applications in Water Resources*, ed. H. Tourno, pp. 1218-1228. New York: ASCE.

Shogan, A.W. 1982. Modular Decomposition and Reliability Computation in Stochastic Transportation Networks Having Cutnodes. *Networks* 12 (1982): 255-275.

Tangena, B. H., and Koster, P. K. 1983. Reliability of Drinking Water Supply Systems. In *Proceedings IWSA Specialized Conference on Security in Water Supply*. Brussels: IWSA.

Tillman, F. A., Hwang, C-L., and Kuo, W. 1980. *Optimization of Systems Reliability*. New York: Marcel Dekker.

U.S. General Accounting Office. 1980. *Additional Federal Aid for Urban Water Distribution Systems Should Wait Until Needs are Clearly Established*. Report to the Congress. Washington, D. C. Quoted in O'Day [1982].

Valiant, L. G. 1979. The Complexity of Enumeration and Reliability Problems. *SIAM J. Comput.* 8 (1979): 410-421.Rosenthal, A. 1981 Series-Parallel Reduction for Difficult Measures of Network Reliability. *Networks* 11 (1981): 323-334.

Walski, T. M. 1984a. *Analysis of Water Distribution Systems*. New York: Van Nostrand Reinhold Co.

----------. 1984b. *Application of Procedures for Testing and Evaluating Water Distribution Systems*. Technical Report EL-84-5. U.S. Army Corps of Engineers.

Walski, T. M. and Pelliccia, A. 1982. Economic analysis of water main breaks. *Journal AWWA* (1982): 140-147.

Wilkov, R. S. 1972. Analysis and Design of Reliable Computer Networks. *IEEE Transactions on Communications* COM-20 (1972): 660-678.

Willie, R. R. 1978. Computer-Aided Fault Tree Analysis, ORC 78-14, Operations Research Center, University of California, Berkeley, CA.

----------. 1979. Combinatorial Methods for Reliability Analysis of Flow Networks. In *Mathematical Theory of Power System Reliability Evaluation*, Memorandum No. UCB/ERL M79/67, Electronics Research Laboratory, Berkeley, CA.

Zemel, E. 1982. Polynomial Algorithms for Estimating Network Reliability. *Networks* 12 (1982): 439-452.