Reliability assessment of urban water networks

Michalis Fragiadakis  
*University of Cyprus, Cyprus*

Dimitrios Vamvatsikos  
*National Technical University of Athens, Greece*

Symeon E. Christodoulou  
*University of Cyprus, Cyprus*

**SUMMARY:**
We present a framework for the seismic risk assessment of aging water supply networks in order to evaluate the water networks of the island of Cyprus. The proposed methodology aims at assessing the performance of water pipe networks based on the pipe vulnerabilities that are determined considering survival curves obtained from historical data of damage. The magnitude-distance seismic scenario adopted is consistent with the seismicity of sites mainly controlled by distant and moderate magnitude events. The network reliability is assessed using Graph Theory, while two alternative approaches are compared for calculating the system reliability: the path enumeration method and Monte Carlo simulation. The methodology proposed can estimate the probability that the network fails to provide the desired level of service and thus allows the prioritization of retrofit interventions and capacity-upgrade actions for the existing water pipe networks.

**Keywords:** Water pipe network, Lifelines, Graph Theory, Monte Carlo Simulation, Survival Analysis, Reliability assessment.

1. **INTRODUCTION**

We present an approach for the seismic risk assessment of water pipe networks. The risk assessment of critical infrastructures, such as water pipe networks, is a major problem for the earthquake engineering community. Recent earthquakes have shown that following a strong earthquake, damage of the lifelines may cause a series of problems immediately after the event, while the residential, commercial and industrial activities may be disrupted for a long time, causing severe economic losses. Such losses may be either direct or indirect. Direct losses are related to the cost of repair, while indirect losses usually depend on the way the economy is affected by the disruption of the lifeline. It is evident that the more developed a society is, the more severe indirect losses should be expected.

Water distribution networks deliver the water from its sources to the customers of the network. Being able to assess the reliability of the network against different hazards helps water distribution agencies prioritize their interventions and ensure a minimum reliability level of the network. Therefore, water distribution agencies are required to develop and implement new methods for monitoring, repairing (or replacing) aging infrastructures, as well as modelling deteriorating infrastructure conditions and proactively devising strategies to keep the networks in operation. In essence, water distribution agencies are faced with the increasingly more complex task of intelligently and efficiently assessing (or modelling) the condition of a pipe network while managing the network in ways that maximize its reliability and minimize its operational and management costs. The question that usually arises is whether an organization should repair or replace deteriorating water mains and, in either case, what should the sequence of any such repairs be as part of a long-term network rehabilitation strategy.

The seismic risk assessment of critical infrastructures has been the subject of extensive past research. For example, Nuti *et al.* (2010) propose a methodology for the reliability assessment of electric power, water and road systems, neglecting the interdependence between the networks, while Poljnašek *et al.*
(2012) propose a method for gas and electricity transmission networks considering the increased vulnerability due to interdependency. Cimellaro (2011) proposed a performance index for evaluating the functionality of a road network during extreme events such as earthquakes that can be utilized within the general PEOPLES framework for measuring disaster resilience indices for a community at various scales. Regarding water supply networks, Romero et al. (2010) discuss the possible seismic hazards and present results that correspond to a severe event in California. Moreover, Wang et al. (2010) propose a methodology for the seismic risk assessment and identification of critical links of water supply systems. The U.S. Federal Emergency Management Agency (FEMA) has also developed the nationally applicable standardized methodology and software program HAZUS-MH MR3 (2003), which estimates potential losses from earthquakes, hurricane winds, and floods.

The reliability of a water pipe network can be calculated if the vulnerability (also termed fragility) of every element of the water network is known. Although water pipe networks consist of several elements (pipes, house connections, tanks, pumps, etc.), focus is given on the pipes, which are, not only the most important component in a piping network but they are also the most difficult component to inspect and replace. Many possible risk-of-failure parameters can be identified (Romero et al. 2010). Our methodology takes into consideration the fragility that corresponds to pipe failures that occur frequently during the everyday operation of the water network and also more severe, but less often, failures due to earthquakes. The pipe vulnerability due to nonseismic causes is assessed using survival analysis techniques on available measurements. Survival analysis considers a number of parameters, e.g. number of observed previous breaks, pipe material, diameter and age that affect the pipe survival curves (Christodoulou 2011) to develop survival/hazard rates and time-to-failure curves for system components based on a multitude of risk-of-failure factors and data stratifications. To account for the vulnerability due to seismic hazard, we rely on the procedure described in the ALA (2001) guidelines and we propose a rational approach for combining it with survival analysis results to include the effect of aging and of the number of previous observed breaks in the network.

Even though the procedure proposed herein is general in scope, it is presented based on performance data from urban water networks in the island of Cyprus. Being a South European island, Cyprus has suffered during the last years from low rainfalls and shortage in its water reserves. Under such conditions, a common practice followed by water distribution agencies is to periodically interrupt the flow in different areas of the city network. This practice offers a more rational treatment of the water resources, but is also considered responsible for an increased failure rate of the pipes of the water network. The worsening failure rate in the water pipe networks of all major cities of the island prompted the initiation of an extensive program of monitoring and keeping track of every damage incident in order to be able to assess the network conditions and assist its proper maintenance. The post-processing of the vast amount of available data is performed using survival analysis tools to produce pipe survival curves. Due to the deterioration of the pipe properties, the seismic vulnerability of these networks increases with time to considerable levels, although the island is located in a moderate seismicity environment.

3. RELIABILITY OF WATER PIPES

3.1. Survival analysis

Survival analysis is a branch of statistics dealing with deterioration and failure over time and involves the modelling of the elapsed time between an initiating event and a terminal event (Hintze 2006). In the case of piping networks such initiating events can be the installation of a pipe, a water-leak observation or the start of a pipe treatment, and cases of terminal events can be a relapse of a previous leak, a fix or a failure. The method is based on estimating the reliability of a system and its lifetime subject to multiple risk factors, and aims to provide answers on the population fraction that survives past an expected lifetime, the effect of the various risk factors on the system’s lifetime and on the probability of survival, and the expected mean time to failure (Hintze 2006, Hosmer et al. 2008). The data values used in the analysis are a mixture of complete and censored observations. In the former
case a terminal event is thought to have occurred, while in the latter case a terminal event has not occurred. A terminal event is assumed to occur just once for each subject.

If \( T \) denotes the elapsed time until the occurrence of a pipe failure, the pipe survival function \( S \) is given by the expression:

\[
S(t) = \int_t^\infty f(x)dx = 1 - F(t)
\]  

(3.1)

Thus the survival function is the probability that the time to failure is longer than some specified time \( t \). Moreover, \( F(t) \) is the cumulative distribution function that denotes the probability that a pipe survives until time \( t \) and \( f(t) \) is the corresponding probability density function. The rate of the survival function is denoted as \( h(T) \) and provides the probability that a pipe at time \( T \) experiences the event in the next instant. The cumulative hazard function \( H(T) \) is the integral of \( h(t) \) from 0 to \( T \). Thus,

\[
S(t) = \exp\left[-\int_t^\infty h(x)dx\right] = \exp\left[-H(T)\right]
\]  

(3.2)

and

\[
h(T) = f(T)/S(T)
\]  

(3.3)

The survival function is usually the primary quantity of interest and is numerically calculated using kernels. In terms of piping networks, the survival function has been shown to be dependent on several factors, among the most important of which being the number of observed previous breaks (NOPB), the age and the material of the material of the pipes (Christodoulou 2011).

### 3.2 Vulnerability analysis of water pipes

The seismic vulnerability of buried pipelines is discussed in the ALA (2001) document. ALA (2001) provides vulnerability curves for water pipes, using observations from past disruptive earthquakes. The failure parameters that affect buried pipes are identified and vulnerability functions are proposed. The vulnerability functions are related to the peak ground velocity (PGV) and the permanent ground deformation (PGD). PGV is related with strong ground shaking caused by seismic wave propagation, while PGD is used to measure factors that include landslides, liquefaction, ground settlement and fault crossing. Other parameters identified are the diameter, the age, the year of construction and discontinuities along the pipe. The recommended pipe vulnerability functions provide the repair rate (RR) per 1000 feet of pipe and have the form:

\[
RR_{PGV} = K_1 \cdot 0.00187 \cdot PGV
\]

\[
RR_{PGD} = K_2 \cdot 1.06 \cdot PGD^{0.319}
\]  

(3.4)

The lognormal standard deviation of the vulnerability functions is 1.15 and 0.74, respectively. The units for PGV and PGD are in/s and in, respectively. Tabulated values are provided for \( K_1 \) and \( K_2 \) depending on the material of the pipe. \( K_i=K_2=1 \) is refers to pipes made from cast iron or asbestos cement. The pipe repairs of Eqn. (3.4) can be due to a complete fracture, a leak or damage to an appurtenance of the pipe, or any other reason that requires the water agency to intervene. For typical water pipe networks, a rule of thumb is that for failure due to wave propagation, 15-20% of failures are breaks and the rest are leaks, while for failures due to PGD 80-85% are breaks that result to the loss of pipeline hydraulic continuity (Ballantyne 1990).

Once the repair rate (RR) is known, i.e. the number of leaks/breaks per pipe length, the Poisson distribution is used to obtain the failure probability of a pipe. The failure probability of a pipe is equal to one minus the probability of zero breaks along the pipe. Using the well-known exponential distribution CDF formula the pipe failure probability \( P_i \) is therefore calculated as (ALA 2001):
\[ P_f = 1 - e^{-RR_L} \]  
(3.5)

where \( RR = \max(RR_{PGV}, RR_{PGD}) \). Note that a Poisson process is a “memoryless” process, and thus Eqn. (3.5) is valid regardless of any previous failures having occurred along the pipe.

### 3.3. Proposed strategy for pipe vulnerability functions

Our goal is to propose a seismic vulnerability assessment methodology for water pipe networks, while exploiting available data of past network failures due to sources other than seismic. Previous research has shown that survival analysis is a valuable tool for implementing methods for monitoring, repairing or replacing aging infrastructures and proactively devising strategies to keep the network in operation. Compared to failures caused by earthquakes, failures from non-seismic causes are more frequent and well distributed in time, while failures due to seismic effects occur intermittently and only when a major earthquake strikes. Thus, it is convenient to compile the data from the two failure causes. This approach is also close to the current practice, since usually it is the water agencies that maintain records of the everyday failure causes, while the seismic effects on the lifelines are usually given a more high-level attention by the civil protection agencies. Moreover, the approaches followed for seismic and non-seismic effects have distinct differences and therefore it is not always straightforward to post-process the data in a manner that allows us to combine consistently pipe survival curves and vulnerability curves.

**Figure 3.1.** Typical survival curves for water pipes, stratified by the NOPB risk-of-failure factor

In our study we combine the vulnerability curves suggested in the ALA (2001) guidelines with available survival curves that were compiled using data available from the water network of the city of Limassol, Cyprus. To this cause we adopt a simplified engineering approach that allows us to quickly combine data that are not similar. Having at our disposal the pipe survival curves (Figure 3.1) of \( S(t) \) versus time (Eqn. 3.1) we know for every time instant (e.g. 1 year) the survival probability of a pipe, depending on the number of previous breaks (NOPB) and the pipe type (e.g. material, diameter, etc). We thus penalize the pipe vulnerability function of Eqn. (3.4) by the ratio of the survival curve of the damaged case (NOPB>0) over the undamaged pipe (NOPB=0). Therefore, after \( t_0 \) days, we define \( k(t_0) = S_a(t_0)/S_{na}(t_0) \geq 1 \), where the subscripts “D” and “UD” stand for “damaged” and “undamaged”, respectively. The modified pipe failure probability that now includes memory of past failures is calculated as:

\[ P_f(t_0) = 1 - e^{-k(t_0)RR_L} \]  
(3.6)

The remaining parameters of Eqn. (3.6) are similar to those of Eqn. (3.5), while now the failure probability depends also on the time instant \( t_0 \) that it is evaluated.
4. RELIABILITY ASSESSMENT OF THE WATER SUPPLY SYSTEM

Once the failure probability of every pipe is known, the performance of the network and its failure probability can be assessed. Depending on the problem at hand, different approaches can be preferred. Perhaps, the most significant parameter that affects the selection of the strategy to follow is how the network performance is measured and thus how the failure probability of the network is defined. In the simplest case, the network fails when it is not able to deliver water from its sources (inflow vertices) to every house connection (outflow vertices). Another, approach would consider the number of customers that are left without water. If such, rather simplified, network performance definitions are adopted, the performance of the network can be quickly evaluated using methods based on Graph Theory (Gibbons 1985). Alternatively, if the failure is defined with respect to hydraulic quantities, i.e., the hydraulic head in every house connection should not be less than a given minimum value, then hydraulic analysis of the network is required. Appropriate software is usually necessary in this case.

In our study we consider as failure of the network its inability to provide water to a consumer/house connection. Therefore, we define the failure probability as the probability of the network being unable to provide water from an inflow source vertex \( i \) to an outflow (e.g. house connection) vertex \( j \). Therefore, if the failure probability to deliver water between \( i \) and \( j \) is \( P_{t,ij} \), the network reliability \( R_{ij} \) is:

\[
R_{ij} = 1 - P_{t,ij}
\]  

(4.1)

For water networks with more than one inflow sources, we consider as failure probability the smallest probability of all paths between the possible sources and the vertex examined. To calculate the network failure probability we propose and compare two alternative approaches. The first is based on the path enumeration method (Henley & Williams 1973) and the second on Monte Carlo simulation.

4.1. Graph theory

Graph theory is a powerful mathematical tool able to describe and manipulate large network data sets. A graph, \( G=(V, E) \) has two basic elements: the nodes (or vertices) denoted by \( V \) and the edges (or links) denoted by \( E \). \( E \) is a set that consists of pairs \((u, v)\) of distinct elements from \( V \) (i.e., \( E \in V \times V \)). Graphs can be directed (arcs) or undirected (edges), depending on whether the movement is possible in one direction or both directions, respectively. Mixed graphs are also possible if they contain both directed and undirected edges. The water pipe networks here examined are directed, since the flow is possible towards one direction only. A directed graph that has a numeric value (i.e. weight) associated with each edge is a weighted directed graph (Gibbons 1985).

![Figure 4.1. Example of a simple graph and its adjacency matrix.](image)

Each graph can be represented either through a graphical shape or with its adjacency matrix (Fig. 4.1). The adjacency matrix of a graph is a matrix with rows and columns labelled by graph vertices, with \( a_{ij} \) equal to 1 or 0 depending on whether \( i \) is adjacent to \( j \) or not. For a simple graph with no self-loops, the adjacency matrix must have zero on the diagonal. Fig. 4.1 illustrates a mixed graph and its adjacency matrix. For a weighted graph, its adjacency matrix could have in position \( a_{ij} \), the weight \( w_{ij} \) assigned to the edge connecting \( i \) and \( j \), which would become zero if there is no connection.

4.2. Reliability bounds using the path enumeration method

The path enumeration method involves identifying all minimal paths of the network. A minimal path
is defined as a set of links so that if these components are all reliable the system functions. A path \( F \) is minimal if it has no subpaths and thus each vertex appears in the path once. If the probability of failure of nodes and links is known, the reliability can be calculated using various techniques (e.g. Ahmad 1988). Once the minimal paths between vertices \( i \) and \( j \) have been identified, the network reliability between \( i \) and \( j \) is calculated as the probability that at least one path is successful. If \( M \) is the number of minimal paths of the network and \( F_m \) is the \( m \)-th minimal path, the exact calculation of the failure probability \( P_{f_{ij}} \) is (Melchers 1999):

\[
P_{f_{ij}} = P(F_i) + P(F_2) - P(F_i \cap F_2)
+ P(F_3) - P(F_i \cap F_3) - P(F_2 \cap F_3) + P(F_1 \cap F_2 \cap F_3)
+ P(F_4) - P(F_i \cap F_4) - P(F_2 \cap F_4) - P(F_3 \cap F_4) - P(F_1 \cap F_2 \cap F_3)
+ P(F_1 \cap F_2 \cap F_4) + P(F_2 \cap F_3 \cap F_4) + P(F_1 \cap F_2 \cap F_3 \cap F_4)
+ P(F_5) - \ldots
= \sum_{k=1}^{M} P(F_k) - \sum_{k=1}^{M} P(F_k \cap F_i) + \sum_{k=1}^{M} \sum_{l<k}^{M} P(F_k \cap F_l \cap F_i) - \ldots (4.2)
\]

where \( \cap \) denotes intersection. However, the size of Eq. (4.2) grows exponentially as \( M \) increases, since \( 2^M-1 \) terms need to be calculated. For example for a network with 10 paths the probability is the sum of \( 2^{10}-1=1023 \) terms, proving that the exact calculation becomes extremely time-consuming even for small networks. To overcome this problem, first- or second-order bounds of the failure probability can be calculated instead. Thus, the bounds of the failure probability can be determined as (Dietlevsen 1979):

\[
P(F_i) + \sum_{k=2}^{M} \max \left[ P(F_k) - \sum_{l=1}^{k-1} P(F_k \cap F_l), 0 \right] \leq P_{f_{ij}} \leq \sum_{k=1}^{M} P(F_k) - \sum_{k=2}^{M} \max \{ P(F_k \cap F_i) \} (4.3)
\]

The above equation provides estimates of the \( P_{f_{ij}} \) which are of sufficient accuracy for engineering purposes. Attention should be paid to the fact that the bounds of Eq. (4.3) depend on the ordering of the \( M \) paths. Approaches for optimal numbering and best results are discussed by Ramachadran (2004) and Song & Der Kiureghian (2003). Instead of second-order bounds, formulas that use third-order bounds also exist in the literature (Zhang 1993), but these also depend on the path numbering. Compared to Eq. (4.3), the variability in this case is much smaller, but the computing effort to calculate the three-term joint probabilities is significant. Another possible shortcoming of this approach is that as the size of the network becomes large, the number of minimal paths may be too large to be accurately identified, stored and processed.

### 4.3. Monte Carlo simulation

When Monte Carlo simulation (MCS) is adopted, the algorithm is based on reducing the network topology. Successful applications of MCS can be found in Javanbarg et al. (2009) and Ramirez-Marquez and Coit (2005). Monte Carlo is a simulation method that relies on performing a large number of simulations. For every simulation a state vector is produced. In this vector, two states can be considered for every pipe: 0-state, which refers to a failed state with probability of \( P_f(t_0) \) (Eqn. (3.6)) and 1-state that corresponds to non-failure with probability \( 1-P_f(t_0) \). Once a state vector is obtained, the failed pipes are removed from the network. Using common graph algorithms we can determine whether a path between vertices \( i \) and \( j \) exists, thus allowing to deliver water from \( i \) to \( j \). If at least one path exists the simulation is successful, otherwise it has failed. The network reliability can then be evaluated by dividing the number of successes with the total number of simulations performed. MCS can easily accommodate both pipe and node failures, while its main shortcoming is its increased computing cost when small probabilities are sought.
5. CASE STUDY

We study a simple water distribution network (Fig. 5.1) which is considered representative of a typical, smaller-scale district metered area (DMA) of the water network of a city in Cyprus. The network has three inputs that could be thought as water reservoirs or inflow pipes and six main valves. In total, the water network has 17 water distribution mains for a total length of 797 pipe segments. The latter number is deduced by summing up the lengths of the arcs in the network (each arc length represents the number of pipe segments on that arc) and corresponds to the total pipe length of the network. We assume unidirectional nodal connections (an acyclic graph) so as to emulate real-life gravity-based water distribution networks. The pipe properties (e.g., material, diameter) are assumed similar throughout the network.

The seismic hazard in the island of Cyprus is mainly controlled by distant and moderate magnitude events. Moreover, city-sized water pipe networks are usually divided to DMAs and each is studied separately. Therefore, it is valid to assume that the seismic intensity will be uniform throughout the DMA water network studied. To estimate the peak ground velocity (PGV) we use the attenuation relationship proposed by Rupakhetty et al. (2011). Thus, PGV is obtained as:

\[
\log(PGV) = -5.17 + 1.98 \cdot M_w - 0.14 \cdot M_w^2 - 0.10 \cdot \log(D^2 + 0.562)
\]

(5.1)

where \(M_w=\text{min}(M_w, M_{sat})\) and \(M_{sat}=7.0\). Care should be given to the fact that PGV is given in cm/s while Eqn. (3.4) is in inches/cm and also on the fact that base 10 logarithms are used in Eqn. (5.1). For this limited example, an array of scenario events is assumed by assigning deterministic values to magnitude \(M_w\) and distance \(D\) in the range 5≤7 and 5≤50km, respectively. We choose to consider as seismic hazard only the wave propagation effects, thus neglecting the PGD term of Eqn. (3.4). The factor that relates the number of previous breaks (NOPB) of survival analysis, \(k(t_0)\) is assumed equal to 1, thus NOPB is equal to 0. The \(K_1\) factor that depends on the pipe material, was assumed equal to 0.70.

Table 5.1. Comparison of the Path Enumeration and Monte Carlo method

<table>
<thead>
<tr>
<th>(M_w)</th>
<th>(D=20)km</th>
<th>(M_w=7.0, D=20)km</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Path Enumeration</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>(R_{ij}^{1.10})</td>
<td>0.98440</td>
<td>0.99926</td>
</tr>
<tr>
<td>(R_{ij}^{2.10})</td>
<td>0.98480</td>
<td>0.99050</td>
</tr>
<tr>
<td>(R_{ij}^{5.10})</td>
<td>0.99090</td>
<td>0.94080</td>
</tr>
<tr>
<td>(R_{ex})</td>
<td>0.98440</td>
<td>0.94080</td>
</tr>
</tbody>
</table>

In Table 5.1, we isolate two \(M_w\), \(D\) cases and we compare the path enumeration method with the Monte Carlo method. For the Monte Carlo we performed \(10^4\) simulations. Since the network has three inflow and one outflow node, we calculate the system reliability \(R_{ij}\) for three combinations of inflow
and outflow nodes. The total system reliability $R_{\text{tot}}$ is considered equal to the minimum of the reliability of the three paths and is also shown in the bottom row of Table 5.1. In general, the differences between the two methods are small, less than 1%, apart from $R_{1\rightarrow10}$ case where the error is 5% and 16% for the $M_w$ 5 and 7 earthquakes, respectively. Since the estimates of the path enumeration method are approximate, we consider as exact solution the one obtained with the Monte Carlo method which was shown to be quite stable when a large number (e.g. $10^7$) simulations are performed. However, in terms of computing cost the path enumeration method was considerably faster.

Fig. 5.2 shows the fragility surfaces of the network, assuming that NOPB=0 ($k(t_0)=1$). The left plot corresponds to the reliability of the 1-10 path and the right is the total system reliability. All results have been obtained with the Monte Carlo method. It is evident that the system becomes less reliable as the seismic event occurs closer to the network and as the earthquake magnitude increases. For comparison we show also the case of $k(t_0)=2$ (Fig. 5.3). This case approximately corresponds to a water network in Cyprus after one year of operation, provided that a small number of previous breaks have occurred. Comparing the figures it is evident that the system quickly deteriorates as the $k(t_0)$ parameter increases. The effect is larger on the more severe $M_w$-$D$ combinations, e.g. compare $(M_w,D) = (7,5)$ with $(M_w, D) = (7,50)$.

6. CONCLUSIONS

A general-purpose methodology for the reliability assessment of water pipe networks has been
presented. The proposed methodology efficiently combines information on failures caused by seismic and non-seismic sources. The more frequent non-seismic failures are typically repaired immediately after the damage is reported to the water agency and result to increasing the vulnerability of the damaged pipe. This sort of information is often available from water agencies and can be postprocessed to provide the pipe survival curves. The seismic vulnerability can be calculated according the methodology of the ALA (2001) guidelines which is straightforward and transparent. Once the pipe failure probabilities are known the reliability of the whole water network can be calculated. We propose and compare two alternative methods for the network reliability. The path enumeration method relies on calculating bounds of the system probabilities and its calculations are very fast but less accurate compared to the time-consuming Monte Carlo method. The proposed approach is a rational, easy to apply and will greatly assist taking proactive measures for the seismic risk mitigation of lifelines.

REFERENCES


Wang, Y. Au, S.-K., Fu, Q. (2010). Seismic Risk Assessment and Mitigation of Water Supply Systems,
Earthquake Spectra, 26:1, 257-273.