Smart Structures and Systems, *Vol. 26, No. 2 (2020) 175-184*
DOI: https://doi.org/10.12989/sss.2020.26.2.175

Accelerated Monte Carlo analysis of flow-based system reliability through artificial neural network-based surrogate models

Sungsik Yoon [1a], Young-Joo Lee [2b] and Hyung-Jo Jung[*1]

[1] *Department of Civil and Environmental Engineering, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea*
[2] *School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Eonyang-eup, Ulju-gun, Ulsan 44919, Republic of Korea*

(*Received October 4, 2019, Revised February 19, 2020, Accepted April 19, 2020*)

Abstract. Conventional Monte Carlo simulation-based methods for seismic risk assessment of water networks often require excessive computational time costs due to the hydraulic analysis. In this study, an Artificial Neural Network-based surrogate model was proposed to efficiently evaluate the flow-based system reliability of water distribution networks. The surrogate model was constructed with appropriate training parameters through trial-and-error procedures. Furthermore, a deep neural network with hidden layers and neurons was composed for the high-dimensional network. For network training, the input of the neural network was defined as the damage states of the k-dimensional network facilities, and the output was defined as the network system performance. To generate training data, random sampling was performed between earthquake magnitudes of 5.0 and 7.5, and hydraulic analyses were conducted to evaluate network performance. For a hydraulic simulation, EPANET-based MATLAB code was developed, and a pressure-driven analysis approach was adopted to represent an unsteady-state network. To demonstrate the constructed surrogate model, the actual water distribution network of A-city, South Korea, was adopted, and the network map was reconstructed from the geographic information system data. The surrogate model was able to predict network performance within a 3% relative error at trained epicenters in drastically reduced time. In addition, the accuracy of the surrogate model was estimated to within 3% relative error (5% for network performance lower than 0.2) at different epicenters to verify the robustness of the epicenter location. Therefore, it is concluded that ANN-based surrogate model can be utilized as an alternative model for efficient seismic risk assessment to within 5% of relative error.

Keywords: Aartificial Neural Networks; surrogate model; accelerated Monte Carlo simulation; seismic risk assessment; flow-based system reliability

1. Introduction

Water distribution networks provide purified water from reservoirs to commercial facilities, industrial facilities, and locals through complex pipeline networks. In particular, lifeline infrastructure, such as water network, are concentrated in the centers of metropolitan cities, which can cause direct and indirect damage to other lifeline facilities (Lee *et al*. 2009). Therefore, when external disturbances such as natural disasters occur, a prompt response is needed to minimize economic damage and casualties (Cerchiello *et al*. 2018, Kim *et al*. 2018).

Research reports from the Pan-American Health Organization (PAHO) have shown that earthquakes have the greatest impact on water network systems among other natural disasters (PAHO 2002). The impact on the water network depends on the frequency and magnitude of the earthquake, but once an earthquake occurs, it can cause serious social disruptions (Jeon and O'Rourke 2005). Therefore, it is important to conduct a seismic risk assessment of the water distribution network at the system level and to establish repair priorities (plans) and recovery strategies.

Various studies have been conducted to assess the seismic risk assessment of lifeline infrastructures such as water networks, power networks, gas networks and transport networks. Early studies adopted connectivity-based approaches to assess system reliability. For example, Esposito *et al*. (2015) conducted a simulation-based seismic risk assessment of a gas distribution network in L'Aquila, Italy, considering the gas network facilities. Rokneddin *et al*. (2013) evaluated bridge network system reliability through an Origin-Destination (O-D) connectivity analysis based on Markov Chain Monte Carlo (MCMC) simulations. Moreover, Dueñas-Osorio *et al*. (2007) performed a seismic hazard analysis of electric power networks considering the interdependency of water networks. In the case of water transmission networks, Yoon *et al*. (2018) proposed a comprehensive framework for seismic risk assessment, and the proposed model was verified using an actual water network located in South Korea.

However, connectivity-based system reliability does not reflect physical environmental conditions such as the nodal

*Corresponding author, Ph.D., Professor,
E-mail: hjung@kaist.edu
[a] Ph.D. Student, E-mail: yss3366@kaist.ac.kr
[b] Associate Professor, E-mail: ylee@unist.ac.kr

ISSN: 1738-1584 (Print), 1738-1991 (Online)

capacity, elevation, and demand. To overcome the limitations of connectivity-based approaches, flow-based system reliability methodologies were proposed. Lee *et al*. (2011) performed a post-hazard flow capacity analysis of a transportation network considering the deterioration of bridges, and Kang *et al*. (2017) conducted the flow analysis of a water pipeline network under seismic conditions. Moreover, Shi and O'Rourke (2006), Wang and O'Rourke (2006), and Bonneau and O'Rourke (2009) developed the GIRAFFE program to measure seismic performance in Los Angeles, California, in extreme events such as earthquakes. Although flow-based network analysis enables more accurate network performance estimation than connectivity-based analysis, it often requires more computational time because hydraulic simulations are required.

For the flow-based seismic risk assessment of a water network, the Monte Carlo simulation (MCS) is used frequently. In the conventional MCS-based system reliability approach, the damaged network is simulated by numerical modeling and removal of impaired network components. The network performance is then estimated by comparing the total and available flow rates of nodal demands at the transmission and sink nodes. However, conventional MCS methods require significant computation time for hydraulic simulations, and the time cost increases exponentially as the system dimension increases. In particular, excessive computation time is needed if iterative systems such as optimization problems, resilience estimation, and seismic risk assessment using probabilistic seismic hazard analysis (PSHA) are required.

To reduce the computational time cost through efficient risk assessment, numerous researchers have performed surrogate-based modeling to enable seismic performance assessments (Mangalathu *et al*. 2018, Seo *et al*. 2012, Seo and Linzell 2013). Stern *et al*. (2017) evaluated the MCS-based seismic performance of the California water network using logistic regression and kernel support vector machine classifiers. Dueñas-Osorio and Rojo (2011) proposed radial topology reliability assessment tools and explored customer service availability and errors. Moreover, Kang *et al*. (2008) and Lim and Song (2012) proposed non-simulation based the matrix-based system reliability (MSR) method and recursive decomposition algorithm (RDA) for the efficient risk assessment of lifeline networks, respectively. However, their research was only applied to connectivity-based network systems and was demonstrated in small networks with system dimensions of less than 100. In particular, previous studies did not focus on reducing the computation time cost and did not consider flow-based system reliability. Therefore, previous studies are not suitable for flow-based system reliability assessment of large dimensional networks for accelerated MCS.

In view of the above, this paper proposes an Artificial Neural Network (ANN)-based surrogate model for accelerated MCS of flow-based system reliability. In particular, the construction of the ANN-based surrogate model is described, and the accuracy and computation time of the surrogate model are compared with the conventional direct method. For flow-based network simulations, an EPANET-based MATLAB computer code was developed to implement the pressure-driven analysis.

To demonstrate the ANN-based surrogate model, an actual water distribution network located in South Korea was adopted, and the network map was reconstructed based on geographic information system (GIS) data. The ANN-based surrogate model was randomly sampled in the range of earthquake magnitudes of 5.0–7.5, and the network damage status and system performance were utilized as training data. To verify the performance prediction of the surrogate model, the prediction results were compared with the results of the conventional direct calculation method in the trained epicenter, and different epicenter locations were adopted to confirm their robustness.

The remainder of this paper is organized as follows. In Section 2, theoretical backgrounds are introduced including the flow-based system reliability analysis, and MCS framework for direct calculation and ANN-based surrogate model are discussed in Section 3. From the proposed surrogate model, a numerical example is demonstrated in Section 4. Finally, Section 5 summarizes the key findings of this research and provides suggestions for future research direction.

2. Flow-based system reliability

2.1 Flow-based network analysis

Network analysis methods are classified into connectivity-based and flow-based methods. Connectivity-based network analysis determines network performance based on failure conditions of nodes or links, regardless of facility capacity and demand. Therefore, the network performance is focused on the connection between the source and the sink node, and does not reflect the physical conditions of the actual flow. However, an actual water distribution network must consider not only the elevation of the target area but also environmental conditions such as the pipe diameter, node demand, and pressure. In addition, as the flow amount can be controlled according to the failure states of the link (pipeline), the sink node performance should be determined to reflect the states of all transmission nodes and links (Guidotti *et al*. 2016, Yoon *et al*. 2020).

In this study, the hydraulic analysis tool EPANET, which was developed by the US Environmental Protection Agency (EPA), was utilized to evaluate the exact network performance of water distribution networks. The traditional EPANET program employs a demand-driven analysis (DDA)-based solver to calculate the nodal pressure. However, under unsteady state conditions, non physical values such as negative pressure can be measured at the water facility (edge or node). Therefore, in this study, EPANET-based MATLAB code that enables a hydraulic simulation of damaged networks with a pressure-based analysis (PDA) approach is implemented.

2.2 Ground motion prediction equation

An earthquake is a phenomenon in which the energy released between different faults is transmitted to the ground surface. Various ground motions can be generated depending on the path of the seismic wave propagation, geotechnical environment, and local site characteristics. Thus, a mathematically simple expression called the ground motion prediction equation (GMPE) is utilized to consider these complex physical phenomena (Ambraseys *et al.* 2005a, b, Boore *et al.* 2003, Joyner and Boore 1993). In general, ground motion can be expressed as the probability distribution of conditionally selected seismic intensity measures such as the peak ground acceleration (PGA), peak ground velocity (PGV), peak ground deformation (PGD) depending on the propagation paths or local site, and soil environments (Abrahamson and Youngs 1992, Esposito and Iervolino 2012, Goda and Hong 2008).

In this study, the PGV seismic attenuation law proposed by Wang and Takada (2005) is utilized

$$\begin{aligned} log\left(\overline{PGV_j}\right) &= 0.725M_i + 0.00318H - 0.519 \\ &\quad -1.318\, log\left(R_{ij} + 0.334e^{0.653M_i}\right) \end{aligned} \tag{1}$$

where $\overline{PGV_j}$ represents the mean velocity (m/s) at the *j*-th site, M_i is the earthquake magnitude at the epicenter, H is the length of the focal depth (km), and R_{ij} is the distance between the epicenter and the *j*-th site.

To represent the uncertainty of the generated ground motion, the spatial correlation terms (inter-events and intra-events) of the ground motion intensity are also considered. Inter- and intra-events represent the uncertainty of the ground motion caused by the intrinsic characteristics of the seismic waves, and seismic wave propagation paths and soil conditions, respectively. This is expressed as the following equation (Sokolov *et al.* 2010, Wagener *et al.* 2016)

$$\rho_{Total} = \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\varepsilon^2} + \frac{\sigma_\varepsilon^2}{\sigma_\eta^2 + \sigma_\varepsilon^2}\rho(\Delta_{ij}) \tag{2}$$

where σ_η^2 and σ_ε^2 represent the predefined residual terms of inter- and intra-events, and $\rho(\Delta_{ij})$ represents the spatial correlation of the ground motion intensity. In this study, the following spatial correlation equation proposed by Goda and Hong (2008) was adopted

$$\rho\left(\Delta_{ij}\right) = e^{(-0.509\sqrt{\Delta})} \tag{3}$$

where Δ_{ij} indicates the distance between the *i*-th site and *j*-th site.

2.3 Seismic vulnerability analysis

Once the magnitude of the ground motion is determined, the failure and leakage probability of the pipeline should be evaluated. According to Hazus-MH (FEMA 2003), PGV is known as a suitable parameter for damage caused by strong ground movement, while PGD is known to be suitable for damage from landslides, ground settlement, and liquefaction.

The failure probability of a pipeline is modeled with the repair rate (number of breaks or leakages per unit length of the pipeline) and is proposed in various equation forms based on empirical seismic data. In this study, the equation proposed by O'Rourke and Ayala (1993) was adopted. In addition, modification factors according to the characteristics of the pipes (diameter, materials, soil conditions, etc.) proposed by Isoyama *et al.* (2000) are reflected, and the survival function proposed by Park *et al.* (2010) is utilized to evaluate the deterioration of the pipeline performance against an earthquake, which is represented by the following equation (Yoon *et al.* 2018)

$$repair\ rate\ (RR_i) = \frac{1}{S(t)}C_1C_2\kappa(PGV_i)^\tau \tag{4}$$

where $S(t)$ is the survival probability of the *i*-th pipeline after being buried for t hours, C_1 represents the modification factor according to the material of the pipeline, C_2 represents the modification factor according to the diameter of the pipe, and κ and τ are the scaling and exponent parameters, respectively.

Using the proposed repair rate (RR), the failure probability of a buried pipeline can be represented by the following equation

$$P_{break,i} = 1 - e^{-RR_iL_i} \tag{5}$$

where $P_{break,i}$ and L_i are the breakage probability and length of the *i*-th pipe, respectively. In addition, the leakage failure probability proposed by Okumura and Shinozuka (1991) was introduced to consider the leakage state. In their research, the leakage failure probability of a pipeline is assumed to be five times the breakage failure probability

$$P_{leak,i} = 5 \times P_{break,i} \tag{6}$$

2.4 Numerical modeling of network structure

For the numerical modeling of buried pipelines, damage states can be classified into three types: breakage, leakage, and intact pipe. In this study, the broad method proposed by Hwang *et al.* (1998) was adopted to evaluate the performance of a pipeline network. In the EPANET program, the emitter function is used to represent the leakage and breakage in the pipeline, and the discharge flow can be calculated based on the emitter coefficient, which is expressed as the following equation

$$Q_{dis,i} = Cp_i^{\gamma} \tag{7}$$

where p_i represents the nodal pressure of the *i*-th damaged pipeline, and C and γ represent the emitter coefficient and exponent, respectively. According to a study conducted by Puchovsky (1999), the sprinkler model well predicts the discharge flow when γ is 0.5. In addition, when the above formula is substituted into the orifice formula, the emitter coefficient can be calculated by the following formula

$$C = C_0A\sqrt{2g} \tag{8}$$

where C_0 represents the flow coefficient (empirically, 0.64 is utilized), A represents the cross-sectional area of the damaged pipeline, and g is the gravitational acceleration. In general, the area of breakage in the pipeline is assumed to be 20% of the entire area, and the area of leakage in the pipeline is assumed to be 3% of the entire area

(a) Damaged pipeline (leakage or breakage) (b) Numerical modelling for hydraulic analysis

Fig. 1 Numerical modeling scheme for damaged (leakage or breakage) pipeline structure (Yoon *et al*. 2020)

(Farahmandfar and Piratla 2017).

The hydraulic analysis process consists of calculating the discharge flow (hydraulic analysis I) and calculating the available nodal pressure (hydraulic analysis II). Fig. 1 shows a numerical modeling technique for calculating the discharge flow rate. First, when a pipeline with breakage or leakage is determined by seismic analysis, the discharge flow rate $Q_{dis,i}$ of the i-th pipeline can be estimated by hydraulic analysis I. Once a leak or breakage flow is calculated, it is applied in EPANET by updating it to the base demand $Q_{base,i}$ of the front node in the flow path. Then, the pipeline is closed if there is a breakage in the pipeline. Through the proposed hydraulic modeling process, the discharge flow of a damaged pipeline is considered as a nodal demand, and the extent of damage is controlled by the emitter coefficient.

Hydraulic analysis II is an analysis phase to evaluate the performance of a damaged network. In the network analysis, the performance of a node is evaluated by the pressure, and the desirable pressure for the required demand is expressed using the following equation (Gupta and Bhave 1996)

$$P_{des,i} = P_{min,i} + R_i(Q_{req,i})^m \tag{9}$$

where $P_{min,i}$ indicates the minimum nodal pressure of the entire network (15 m is utilized in this study), R_i is the resistance coefficient (0.1 is utilized), and m is the exponent coefficient (2 is utilized).

The performance of a damaged network can be evaluated by employing the head-outflow relationship (HOR), which represents the relationship between the nodal serviceability and available nodal pressure. In this study, the HOR equation proposed by Wagner *et al*. (1988) is adopted to estimate the nodal serviceability of the i-th node (NS_i) according to the available pressure, which is expressed as follows

$$NS_i = \begin{cases} 0 & if\ P_{avl,i} < P_{min,i} \\ (\frac{P_{avl,i} - P_{min,i}}{P_{des,i} - P_{min,i}})^{1/n} & if\ P_{min,i} < P_{avl,i} < P_{des,i} \\ 1 & if\ P_{avl,i} > P_{des,i} \end{cases} \tag{10}$$

$$Q_{avl,i} = Q_{req,i} \times NS_i \tag{11}$$

where n denotes the serviceability coefficient (2 is utilized), $P_{avl,i}$ denotes the available nodal pressure of the i-th node, NS_i represents the nodal serviceability of the i-th node according to $P_{avl,i}$, and $Q_{avl,i}$ represents the available nodal demand of the i-th node according to the nodal serviceability.

2.5 Performance indicator

When the available nodal demand is calculated across the network through the hydraulic analysis phase, the network performance can be predicted at the system level. Various approaches can be adopted to evaluate the network performance depending on the network analysis (i.e., connectivity or flow analysis). In particular, different performance indicators are available depending on the size, type, and topology of the network. Therefore, it is important to define appropriate performance indicators that can reflect the flow characteristics of the water distribution network.

In this study, the performance indicator utilized was the system serviceability index (S_s) proposed by Wang *et al.* (2010). S_s can be expressed as the ratio of the required nodal demand of the sink node before the earthquake and the available nodal demand after the earthquake

$$S_S = \frac{\sum_{i=1}^{n} Q_{avl,i}}{\sum_{i=1}^{n} Q_{req,i}} \tag{12}$$

where n is the total number of sink nodes.

3. Surrogate model for accelerated Monte Carlo analysis of system reliability

3.1 Conventional Monte Carlo simulation framework for direct calculation

Fig. 2 shows the conventional MCS framework for the performance evaluation of water distribution networks. An EPANET input file must be generated based on GIS data to perform direct calculations using MCS. When the network map is reconstructed, the ground motion can be predicted with spatial correlation, and the failure probability of the network components can be evaluated. The failure probability of the network components is estimated by Hazus-MH, and the network states are determined by random sampling between 0 and 1. If the random number generated in each sampling is less than the probability of failure, then it is considered destroyed; otherwise, the condition is intact.

If damaged network components are identified during the earthquake generation phase, it moves to the hydraulic analysis phase. Because damage to the pipeline cannot be simulated in the EPANET analysis, the hydraulic analysis is utilized to evaluate the discharge flow rate of the pipeline where breakages and leaks have occurred, and then the network is updated with the nodal demands of the front

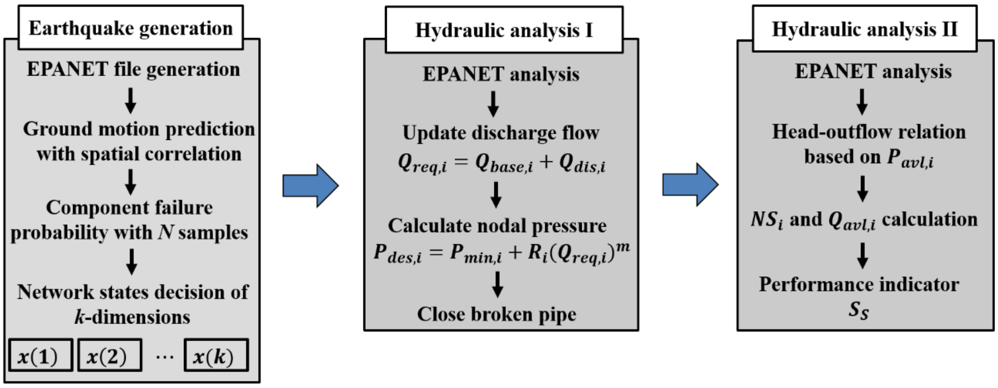

Fig. 2 Conventional MCS framework for direct calculation of network performance

node. Hydraulic analysis I enables the numerical modeling of damaged pipelines under unsteady-state conditions, and the calculated $Q_{req,i}$ and $P_{des,i}$ are moved to hydraulic analysis II.

The phase of hydraulic analysis II is the performance assessment of the damaged water distribution network. When the available nodal pressures are estimated through hydraulic analysis II, the nodal serviceability is determined by comparing them with the desired nodal pressure calculated in the previous phase of hydraulic analysis I. This MCS framework enables the evaluation of the available nodal demands and system serviceability of the sink node at the i-th sampling. The hydraulic analysis performs iterative calculations until the network performance (S_S) is within the convergence range or the MCS reaches N_{max}.

3.2 Proposed ANN-based surrogate model

ANN is an artificial intelligence technique that recognizes patterns of input and output data (hidden layers and neurons) and expresses the relationship between the input and output as a function through supervised learning. Because ANN technology is easy to use and has fast computation, various researchers have adopted the ANN technique in the engineering domain such as structural damage detection (Kim *et al*. 2008, Nguyen *et al*. 2019, Li *et al*. 2015), structural response prediction (Hakim and Razak 2014, Onat and Gul 2018, Shahbazi *et al*. 2014), structural control (Akin and Sahin 2017), and structural monitoring (Rizzo and Lanza 2006). In this study, an ANN-based surrogate model was employed to accelerate the seismic performance evaluation of a water distribution network under seismic conditions.

The MCS-based direct calculation introduced in Section 3.1 requires a significant computation time due to hydraulic analyses (see Fig. 2). Therefore, this study aims to quickly predict the network performance of a water network through an ANN-based surrogate model. In particular, as lifeline structures such as waterworks are buried underground, the network structure remains constant for all systems. Therefore, even though it takes a few hours to build a surrogate model, it can be effectively and consistently utilized for excessively iterative systems such as optimization problems, seismic resilience estimation, and seismic risk analysis employing PSHA.

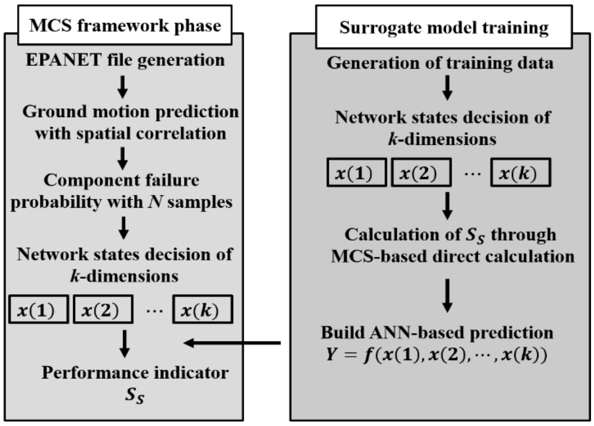

Fig. 3 Proposed framework for approximate system reliability with ANN-based surrogate model

The input data for water network performance prediction are k-dimensional network component states during earthquakes (0: damaged, 1: breakage, 2: leakage), and the output data is defined as system performance between 0 and 1. Fig. 3 shows the system reliability approximation framework using the ANN-based surrogate model. Unlike the conventional direct calculation method, the surrogate model method constructs a response function of the network performance to quickly calculate the system performance when the network conditions are determined. The proposed accelerated framework uses k-dimensional network states and the system serviceability to build predictive models, which can significantly reduce the excessive computation time due to hydraulic analyses. The ANN-based prediction model evaluates the network performance through a built-in surrogate model, which enables accelerate MCS over the traditional MCS framework.

4. Numerical example

4.1 Description of water distribution network

To demonstrate the proposed ANN-based surrogate model, an actual water distribution network located in A

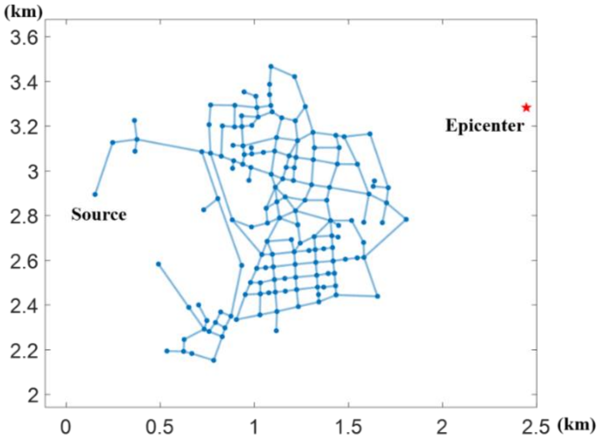

Fig. 4 Reconstructed network map of target water distribution network

city, South Korea, was adopted. GIS source data was provided by the A-city waterworks headquarters, and the network map was reconstructed by post-processing the source data.

The A-city water network supplies purified water from one source node to 66 sink nodes through 85 transmission nodes. The total length of the pipeline is 18.9 km, and the maximum water supply is 9800 m^3/day in an area of 6 km^2. The diameter of the pipeline is distributed between 50 and 400 mm (mainly 150–300 mm). The pipeline type consists of ductile cast iron pipes (DCIP), polyethylene pipes (PE), cast iron pipes (CIP), and polyvinyl chloride pipes (PVC). The deterioration of the pipeline was investigated as at 24.9 years after burial. Moreover, the elevation of the node was estimated using the Google Earth Pro and GIS data, and the elevation of the demonstrated target region was in the range of 19–34 m. Fig. 4 shows a reconstructed network map of the A-city water distribution network.

4.2 Generation of training data

In this section, a surrogate model constructed to predict the seismic performance of the target water network is presented. The epicenter for the generation of training data was selected using the location of a historical earthquake of magnitude 5.4 with a focal depth of 10 km (see Fig. 4). In addition, the magnitude of the earthquake was between 5.0 and 7.5 with a uniform distribution. This was determined by random sampling, and the ANN surrogate model was trained using a total of 300,000 data items. Table 1 shows the ANN network configuration for the surrogate model. To avoid overfitting or underfitting the training and test data, 10,000 sufficient epochs were used and the dataset was set at a ratio of 0.7: 0.15: 0.15 (training data: test data: validation data). In addition, a learning rate of 0.01, momentum constant of 0.9, and error tolerance of 0.0001 were adopted to select the learning parameters. The selected parameters were determined through a trial procedure, and sufficient numbers of hidden layers and neurons were utilized to predict the solution of the 218-dimensional function. In this study, regarding the trade-off between

Table 1 ANN properties for surrogate model

Network dimensions (output)	218 (1)
Network type	Feed-forward back propagation
Training function	Gradient descent with momentum and adaptive learning rate
Adaption learning function	Gradient descent with momentum weight and bias learning function
Performance function	Mean Squared Error (MSE)
Number of layers	15
Number of neurons	15
Transfer function	Tan-Sigmoid transfer function

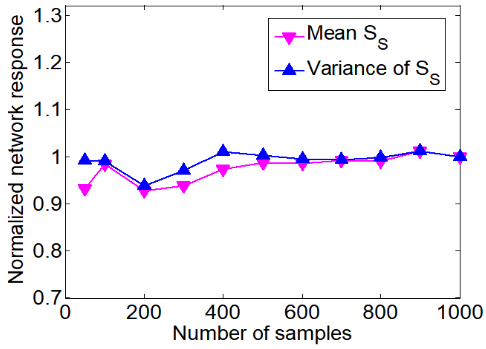

Fig. 5 Convergence of mean and variance of system serviceability according to number of samples

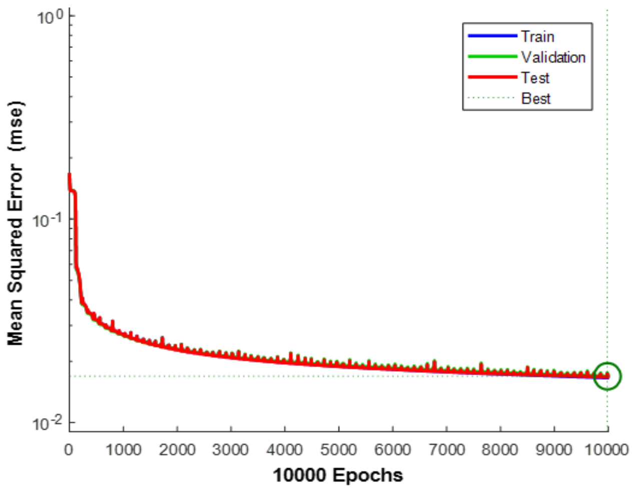

Fig. 6 ANN training performance according to epoch

MCS required to train ANN and the accuracy of ANN, MCS with 500 samples were utilized. Fig. 5 shows the mean and variance of the system services.

Fig. 6 shows the training, test, and validation performance according to the epoch. In this study, an appropriate number of 10,000 epochs was selected to prevent the underfitting and overfitting of training data. As a result of network training, the correlation between the predicted and observed values was identified as follows: training data (0.9687), validation data (0.9865), test data

Table 2 Comparison of results between direct calculation method and ANN-based method

Earthquake magnitude	Direct calculation method		ANN-based method	
	S_S	Computation time (seconds)	S_S	Computation time (seconds)
5.0	0.9667	460s	0.9735	1.18s
5.5	0.9346	475s	0.9279	1.21s
6.0	0.8838	640s	0.8762	1.26s
6.5	0.7768	650s	0.7689	1.29s
7.0	0.4426	675s	0.4368	1.27s
7.5	0.1742	740s	0.1681	1.27s

(0.9677), and all data (0.9685).

4.3 Performance prediction: Trained epicenter

In this section, the performance estimation of trained epicenters conducted using trained surrogate models is presented. The location of the input earthquake was the same as that of the epicenter for generating training data, and six earthquake magnitudes were chosen: 5.0, 5.5, 6.0, 6.5, 7.0, and 7.5. In addition, the results of the ANN-based prediction model (performance estimation and computation time) were compared with the results of the conventional MCS framework through direct calculation.

Table 2 lists the network performance and computational time cost using the MCS framework results based on direct calculation and the ANN-based prediction model for 10,000 samples. The performance of the network tends to decrease as the earthquake magnitude increases. In particular, S_S decreases rapidly when the earthquake magnitude is larger than 6.5. This is consistent with seismic design standards, in which the water distribution network in South Korea was designed to withstand an earthquake magnitude of 5.7 to 6.4. In addition, the results of the direct calculation method and the ANN-based method are found to have a relative error within 0.3 to 3%. When the network performance exceeds 0.5, the relative error is found to be within 1%, but the relative error is more sensitive as S_S decreases below 0.4. For computation time costs, the direct calculation method requires a large computational time (more than 460 seconds) to perform 500 MCS for convergence due to hydraulic analyses, but for the surrogate model, the computation was completed in around 1.2 s using the built-in prediction function.

Table 3 Cumulative ratio of relative errors between direct calculation method and ANN-based method

Relative error		1%	3%	5%	7%	10%
Earthquake magnitude	5.0	0.875	0.996	0.997	0.997	0.999
	5.5	0.857	0.966	0.974	0.988	0.996
	6.0	0.838	0.929	0.944	0.958	0.972
	6.5	0.814	0.915	0.924	0.948	0.974
	7.0	0.806	0.908	0.914	0.938	0.977
	7.5	0.754	0.862	0.897	0.929	0.972

Fig. 7 shows the convergence of network performance when the earthquake magnitude is 7.0 and 7.5. The solid line represents the mean network performance, and the dashed line represents the coefficient of variation (COV) of the network performance. As the number of samples increases, the network performance tends to converge. For an earthquake magnitude of 7.0, S_S is evenly distributed from 0 to 1, so the convergence of network performance is proven only when a sufficient MCS number is obtained (more than 4000 samples). However, with an earthquake magnitude of 7.5, the network performance is distributed below 0.5. Thus, the network performance can be achieved with a small number of samples (more than 2500). Especially for COV, the network performance tends to converge faster at an earthquake magnitude of 7.5.

Table 3 lists the cumulative ratio distribution of the relative errors for S_S between the surrogate model and the direct calculation method. As the magnitude of an earthquake increases (decreases in S_S), the cumulative ratio of relative errors tends to decrease. In particular, if the magnitude of the earthquake is greater than 7.0, then the network performance will be significantly reduced, leading to an increase in the variation of the relative error. This can be attributed to a decrease in the absolute range of the relative error as the magnitude of the earthquake increases (a decrease in network performance). Moreover, it was confirmed that the cumulative ratio of the total number of samples gradually increased as the criterion of the relative error increased.

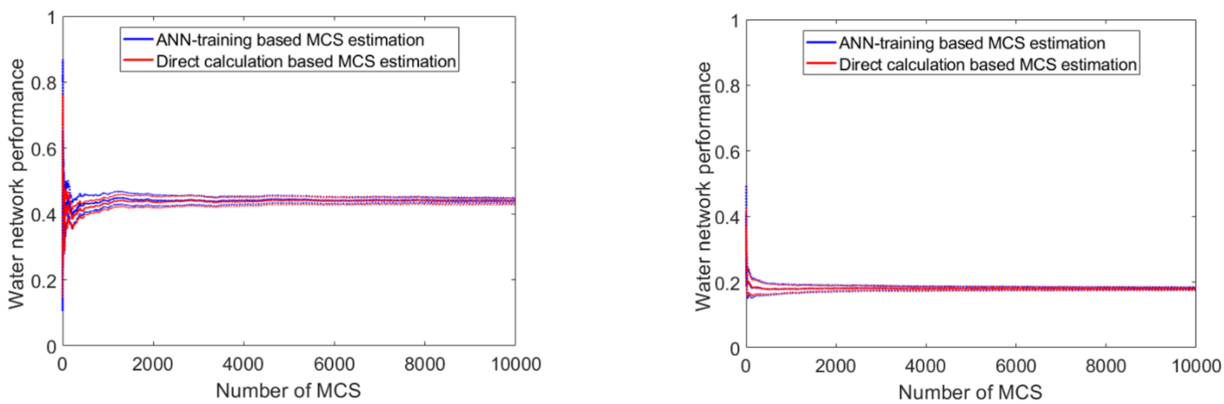

Fig. 7 Convergence of network performance using direct calculation method and ANN-based method

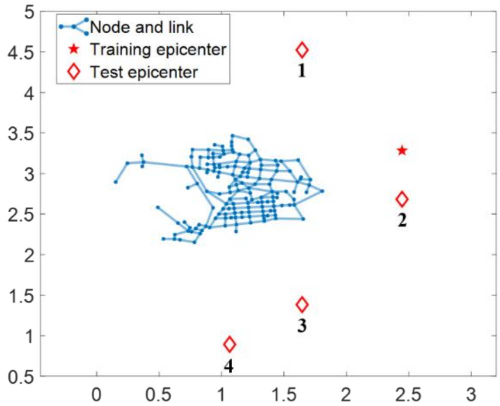

Fig. 8 Location of test earthquake locations for the robustness of epicenters

4.4 Performance prediction: Robustness of epicenters

In Section 4.3, the surrogate model was tested on trained epicenters. However, the epicenter can be located depending on the distribution of faults. Therefore, this study aims to verify the surrogate model by selecting the epicenter location around A-city (near-fault). Fig. 8 shows the location of the test earthquake and the training epicenter for verification of the surrogate model. Six earthquake magnitudes between 5.0 and 7.5 were considered at each location.

Table 4 compares the performance of the ANN-based and direct calculations according to the epicenter location and earthquake magnitude. The network performance estimates of the surrogate models were found to be within 6% of the relative error compared with that of the direct calculation method. In particular, when the earthquake magnitude is less than 6.5, it can be confirmed that the relative error can be predicted within 2%. However, as the magnitude of the earthquake gradually increases, the relative error tends to increase. This is because S_S decreases as the magnitude of the earthquake increases, making it more sensitive to relative errors. The surrogate model also shows that the relative error increases because the network performance is more widely distributed between 0 and 1 when the earthquake magnitude is between approximately 6.5 and 7.0.

Table 4 Comparison of results between direct calculation method and ANN-based method for different epicenters

Test epicenter	System serviceability (S_S)					
Location 1	5.0	5.5	6.0	6.5	7.0	7.5
ANN-based	0.9529	0.9272	0.8732	0.7611	0.5262	0.2307
Direct calculation	0.9466	0.9176	0.8578	0.7461	0.5128	0.2175
Relative error (%)	0.66	1.05	1.79	2.01	2.61	6.07
Location 2	5.0	5.5	6.0	6.5	7.0	7.5
ANN-based	0.9526	0.9084	0.8648	0.7293	0.4907	0.2171
Direct calculation	0.9479	0.9161	0.8538	0.7460	0.4761	0.2057
Relative error (%)	0.49	0.84	1.29	2.24	3.07	5.54
Location 3	5.0	5.5	6.0	6.5	7.0	7.5
ANN-based	0.9512	0.9142	0.8454	0.7293	0.5093	0.2223
Direct calculation	0.9439	0.9206	0.8614	0.7432	0.4960	0.2113
Relative error (%)	0.77	0.69	1.85	1.87	2.68	5.21
Location 4	5.0	5.5	6.0	6.5	7.0	7.5
ANN-based	0.9461	0.9349	0.8810	0.7670	0.5435	0.2338
Direct calculation	0.9516	0.9263	0.8715	0.7778	0.5298	0.2246
Relative error (%)	0.58	0.93	1.09	1.39	2.59	4.09

5. Conclusions

Conventional MCS-based direct calculation methods require excessive computation time due to iterative hydraulic analyses. In this study, an ANN-based surrogate model for accelerated MCS of the seismic performance estimation of water distribution networks was proposed. To train the surrogate model, the network component states (0: intact, 1: breakage, 2: leakage) and the performance indicator (system serviceability) were selected as input and output data, respectively. In the neural network, appropriate training parameters were selected through a trial procedure, and a deep neural network was set up to compute a large network of 218 dimensions with a sufficient number of hidden layers and neurons to avoid overfitting or underfitting.

To evaluate the network performance (exact solution) of the training data, an EPANET-based hydraulic analysis was utilized and the PDA approach was employed to represent the network in an unsteady-state condition. In addition, EPANET-based MATLAB computer code was developed to enable PDA-based hydraulic analysis. To demonstrate the surrogate model, the water distribution network of A-city in South Korea was adopted, and the network map was reconstructed based on GIS data. The epicenter for training data was selected based on historical earthquake data, and the earthquake magnitude was determined by random sampling of earthquake magnitudes between 5.0 and 7.5 with uniform distribution.

To verify the constructed surrogate model, the epicenter was considered as the epicenter of the training data, and the results of the direct calculation method and the surrogate model were compared. The numerical results showed that S_S of the surrogate model was predicted within 3% of the relative error compared to the direct calculation method. For earthquakes above a magnitude of 7.0, the relative error was approximately 3%, while for earthquakes below a magnitude of 7.0, the relative error was less than 1%. This can be attributed to the fact that the relative error increases because the network performance decreases as the earthquake magnitude increases. For the computational time cost, the response function built by the surrogate model showed that the computation time was significantly reduced because there was no hydraulic analysis process. In addition, a numerical analysis was performed using four different historical earthquake data items to verify the robustness of the epicenter location of the surrogate model.

The numerical results showed that the surrogate model accurately predicts the network performance compared to traditional MCS-based direct computation methods and that the relative errors decrease as the network performance increases. Therefore, the proposed model can be used as a database for decision modeling to quickly determine the restoration priority of water supply facilities and to establish a recovery strategy through the spatially correlated ground motion included in the surrogate model

So far, this paper mainly focuses on developing a new framework to construct a surrogate model for accelerated MCS of the seismic performance estimation of water distribution networks. For this reason, the ANN technique was applied for the sake of simplicity. The results showed that the surrogate model using the ANN model showed an accurate prediction of network performance that fared well as compared with conventional direct Monte Carlo analysis. In the future, specialized algorithms, such as a multi-view graph convolutional network, should be applied to efficiently handle the complex graph data of various lifeline facilities, such as water treatment plants, water storage tanks, and water pumping plants.

Acknowledgments

The research described in this paper was financially supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20181510102410). This work was also financially supported by the National Research Foundation Korea (NRF) Grant funded by the Korean government (MSIP) (No. 2017R1A5A1014883).

References

Abrahamson, N.A. and Youngs, R. (1992), "A stable algorithm for regression analyses using the random effects model", *Bull. Seismol. Soc. Am.*, **82**, 505-510.

Akin, O. and Sahin, M. (2017), "Active neuro-adaptive vibration suppression of a smart beam", *Smart Struct. Syst., Int. J.*, **20**(6), 657-668. https://doi.org/10.12989/sss.2017.20.6.657

Ambraseys, N., Douglas, J., Sarma, S. and Smit, P. (2005a), "Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: horizontal peak ground acceleration and spectral acceleration", *Bull. Earthq. Eng.*, **3**, 1-53. https://doi.org/10.1007/s10518-005-0183-0

Ambraseys, N.N., Douglas, J., Sarma, S.K. and Smit, P.M. (2005b), "Equations for the Estimation of Strong Ground Motions from Shallow Crustal Earthquakes Using Data from Europe and the Middle East: Vertical Peak Ground Acceleration and Spectral Acceleration", *Bull. Earthq. Eng.*, **3**, 55-73. https://doi.org/10.1007/s10518-005-0186-x

Bonneau, A.L. and O'Rourke, T.D. (2009), "Water supply performance during earthquakes and extreme events", MCEER Technical Report-MCEER-09-0003.

Boore, D.M., Gibbs, J.F., Joyner, W.B., Tinsley, J.C. and Ponti, D.J. (2003), "Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega Boulevard bridge collapse, West Los Angeles, California", *Bull. Seismol. Soc. Am.*, **93**, 2737-2751. https://doi.org/10.1785/0120020197

Cerchiello, V., Ceresa, P., Monteiro, R. and Komendantova, N. (2018), "Assessment of social vulnerability to seismic hazard in Nablus, Palestine", *Int. J. Disaster Risk Reduct*, **28**, 491-506. https://doi.org/10.1016/j.ijdrr.2017.12.012

Dueñas-Osorio, L. and Rojo, J. (2011), "Reliability assessment of lifeline systems with radial topology", *Comput. Aided Civil Infrastruct. Eng.*, **26**, 111-128. https://doi.org/10.1111/j.1467-8667.2010.00661.x

Dueñas-Osorio, L., Craig, J.I. and Goodno, B.J. (2007), "Seismic response of critical interdependent networks", *Earthq. Eng. Struct. Dyn.*, **36**, 285-306. https://doi.org/10.1002/eqe.626

Esposito, S. and Iervolino, I. (2012), "Spatial correlation of spectral acceleration in European data", *Bull. Seismol. Soc. Am.*, **102**, 2781-2788. https://doi.org/10.1785/0120120068

Esposito, S., Iervolino, I., d'Onofrio, A., Santo, A., Cavalieri, F. and Franchin, P. (2015), "Simulation-based seismic risk assessment of gas distribution networks", *Comput. Aided Civil Infrastruct. Eng.*, **30**, 508-523. https://doi.org/10.1111/mice.12105

Farahmandfar, Z. and Piratla, K.R. (2017), "Comparative evaluation of topological and flow-based seismic resilience metrics for rehabilitation of water pipeline systems", *J. Pipeline Syst. Eng. Pract.*, **9**, 04017027. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000293

FEMA (2003), Multi-Hazard Loss Estimation Methodology Earthquake Model, HAZUS-MH MR3 Technical Manual; United States Department of Homeland Security, Federal Emergency Management Agency, Washington, DC, USA.

Goda, K. and Hong, H.-P. (2008), "Spatial correlation of peak ground motions and response spectra", *Bull. Seismol. Soc. Am.*, **98**, 354-365. https://doi.org/10.1785/0120070078

Guidotti, R., Chmielewski, H., Unnikrishnan, V., Gardoni, P., McAllister, T. and van de Lindt, J. (2016), "Modeling the resilience of critical infrastructure: The role of network dependencies", *Sustain. Resilient Infrastruct.*, **1**, 153-168. https://doi.org/10.1080/23789689.2016.1254999

Gupta, R. and Bhave, P.R. (1996), "Comparison of methods for predicting deficient-network performance", *J. Water Resour. Plan Manag.*, **122**, 214-217. https://doi.org/10.1061/(ASCE)0733-9496(1996)122:3(214)

Hakim, S.,and Razak, H.A. (2014), "Modal parameters based structural damage detection using artificial neural networks-a review", *Smart Struct. Syst., Int. J.*, **14**(2), 159-189. https://doi.org/10.12989/sss.2014.14.2.159

Hwang, H.H., Lin, H. and Shinozuka, M. (1998), "Seismic performance assessment of water delivery systems", *J. Infrastruct. Syst.*, **4**, 118-125. https://doi.org/10.1061/(ASCE)1076-0342(1998)4:3(118)

Isoyama, R., Ishida, E., Yune, K. and Shirozu, T. (2000), "Seismic damage estimation procedure for water supply pipelines", *Proceedings of the 12th World Conference on Earthquake Engineering* (*WCEE*), Auckland, New Zealand, p. 1762.

Jeon, S.-S. and O'Rourke, T.D. (2005), "Northridge earthquake effects on pipelines and residential buildings", *Bull. Seismol. Soc. Am.*, **95**, 294-318.

Joyner, W.B. and Boore, D.M. (1993), "Methods for regression analysis of strong-motion data", *Bull. Seismol. Soc. Am.*, **83**, 469-487.

Kang, W.-H., Song, J. and Gardoni, P. (2008), "Matrix-based system reliability method and applications to bridge networks", *Reliab. Eng. Syst. Saf.*, **93**, 1584-1593. https://doi.org/10.1016/j.ress.2008.02.011

Kang, W.-H., Lee, Y.-J. and Zhang, C. (2017), "Computer-aided analysis of flow in water pipe networks after a seismic event", *Mathe. Probl. Eng.* https://doi.org/10.1155/2017/2017046

Kim, J.-T., Park, J.-H., Koo, K.-Y. and Lee, J.-J. (2008), "Acceleration-based neural networks algorithm for damage detection in structures", *Smart Struct. Syst., Int. J.*, **4**(5), 583-603. https://doi.org/10.12989/sss.2008.4.5.583

Kim, J., Deshmukh, A. and Hastak, M. (2018), "A framework for assessing the resilience of a disaster debris management system", *Int. J. Disaster Risk Reduct.*, **28**, 674-687. https://doi.org/10.1016/j.ijdrr.2018.01.028

Lee, D.H., Kim, B.H., Lee, H. and Kong, J.S. (2009), "Seismic behavior of a buried gas pipeline under earthquake excitations", *Eng. Struct.*, **31**, 1011-1023. https://doi.org/10.1016/j.engstruct.2008.12.012

Lee, Y.-J., Song, J., Gardoni, P. and Lim, H.-W. (2011), "Post-hazard flow capacity of bridge transportation network considering structural deterioration of bridges", *Struct. Infrastruct. Eng.*, **7**, 509-521. https://doi.org/10.1080/15732479.2010.493338

Li, P.H.., Zhu, H.P., Luo, H. and Weng, S. (2015), "Structural damage identification based on genetically trained ANNs in beams", *Smart Struct. Syst., Int. J.*, **15**(1), 227-244. https://doi.org/10.12989/sss.2015.15.1.227

Lim, H.W. and Song, J. (2012), "Efficient risk assessment of lifeline networks under spatially correlated ground motions using selective recursive decomposition algorithm", *Earthq. Eng. Struct. Dyn.*, **41**, 1861-1882. https://doi.org/10.1002/eqe.2162

Mangalathu, S., Heo, G. and Jeon, J.-S. (2018), "Artificial neural network based multi-dimensional fragility development of skewed concrete bridge classes", *Eng. Struct.*, **162**, 166-176. https://doi.org/10.1016/j.engstruct.2018.01.053

Nguyen, D.H., Bui, T.T., De Roeck, G. and Wahab, M.A. (2019), "Damage detection in Ca-Non Bridge using transmissibility and artificial neural networks", *Struct. Eng. Mech., Int. J.*, **71**(2), 175-183. https://doi.org/10.12989/sem.2019.71.2.175

O'Rourke, M. and Ayala, G. (1993), "Pipeline damage due to wave propagation", *J. Geotech. Eng.*, **119**, 1490-1498. https://doi.org/10.1061/(ASCE)0733-9410(1993)119:9(1490)

Okumura, T. and Shinozuka, M. (1991), "Serviceability analysis of Memphis water delivery system", *Proceedings of the 3rd US Conference on Lifeline Earthquake Engineering*, Los Angeles, CA, USA.

Onat, O. and Gul, M. (2018), "Application of artificial neural networks to the prediction of out-of-plane response of infill walls subjected to shake table", *Smart Struct. Syst., Int. J.*, **21**(4), 521-535. https://doi.org/10.12989/sss.2018.21.4.521

PAHO (2002), Emergencies and Disasters in Drinking Water Supply and Sewage Systems: Guidelines for Effective Response; Pan American Health Organization (PAHO).

Park, S., Choi, C.L., Kim, J.H. and Bae, C.H. (2010), "Evaluating the economic residual life of water pipes using the proportional hazards model", *Water Resour. Manag.*, **24**, 3195-3217. https://doi.org/10.1007/s11269-010-9602-3

Puchovsky, M.T. (1999), *Automatic sprinkler systems handbook*, National Fire Protection Association (NFPA).

Rizzo, P. and Lanza, D.S. (2006), "Wavelet-based feature extraction for automatic defect classification in strands by ultrasonic structural monitoring", *Smart Struct. Syst., Int. J.*, **2**(3), 253-274. https://doi.org/10.12989/sss.2006.2.3.253

Rokneddin, K., Ghosh, J., Dueñas-Osorio, L. and Padgett, J.E. (2013), "Bridge retrofit prioritisation for ageing transportation networks subject to seismic hazards", *Struct. Infrastruct. Eng.*, **9**, 1050-1066. https://doi.org/10.1080/15732479.2011.654230

Seo, J. and Linzell, D.G. (2013), "Use of response surface metamodels to generate system level fragilities for existing curved steel bridges", *Eng. Struct.*, **52**, 642-653. https://doi.org/10.1016/j.engstruct.2013.03.023

Seo, J., Dueñas-Osorio, L., Craig, J.I. and Goodno, B.J. (2012), "Metamodel-based regional vulnerability estimate of irregular steel moment-frame structures subjected to earthquake events", *Eng. Struct.*, **45**, 585-597. https://doi.org/10.1016/j.engstruct.2012.07.003

Shahbazi, Y., Delavari, E. and Chenaghlou, M.R. (2014), "Predicting the buckling load of smart multilayer columns using soft computing tools", *Smart Struct. Syst., Int. J.*, **13**(1), 81-98. https://doi.org/10.12989/sss.2013.13.1.081

Shi, P. and O'Rourke, T.D. (2006), "Seismic response modeling of water supply systems", MCEER Technical Report-MCEER-08-0016.

Sokolov, V., Wenzel, F., Jean, W.-Y. and Wen, K.-L. (2010), "Uncertainty and spatial correlation of earthquake ground motion in Taiwan", *Terr. Atmos. Ocean. Sci.*, **21**(6), 905-921. https://doi.org/10.3319/TAO.2010.05.03.01(T)

Stern, R.E., Song, J. and Work, D.B. (2017), "Accelerated Monte Carlo system reliability analysis through machine-learning-based surrogate models of network connectivity", *Reliab. Eng. Syst. Saf.*, **164**, 1-9. https://doi.org/10.1016/j.ress.2017.01.021

Wagner, J.M., Shamir, U. and Marks, D.H. (1988), "Water Distribution Reliability: Simulation Methods", *J. Water Resour. Plan Manag.*, **114**, 276-294. https://doi.org/10.1061/(ASCE)0733-9496(1988)114:3(276)

Wagener, T., Goda, K., Erdik, M., Daniell, J. and Wenzel, F. (2016), "A spatial correlation model of peak ground acceleration and response spectra based on data of the Istanbul Earthquake Rapid Response and Early Warning System", *Soil Dyn. Earthq. Eng.*, **85**, 166-178. https://doi.org/10.1016/j.soildyn.2016.03.016

Wang, Y. and O'Rourke, T.D. (2006), "Seismic performance evaluation of water supply systems", MCEER Technical Report-MCEER-08-0015.

Wang, M. and Takada, T. (2005), "Macrospatial correlation model of seismic ground motions", *Earthq. Spectra*, **21**, 1137-1156. https://doi.org/10.1193/1.2083887

Wang, Y., Au, S.-K. and Fu, Q. (2010), "Seismic risk assessment and mitigation of water supply systems", *Earthq. Spectra*, **26**, 257-274. https://doi.org/10.1193/1.3276900

Yoon, S., Lee, Y.-J. and Jung, H.-J. (2018), "A comprehensive framework for seismic risk assessment of urban water transmission networks", *Int. J. Disaster Risk Reduct.*, **31**, 983-994. https://doi.org/10.1016/j.ijdrr.2018.09.002

Yoon, S., Lee, Y.-J. and Jung, H.-J. (2020), "A comprehensive approach to flow-based seismic risk analysis of water transmission network", *Struct. Eng. Mech., Int. J.*, **73**(3), 339-351. https://doi.org/10.12989/sem.2020.73.3.339

BS