SAND20XX-XXXXR
LDRD PROJECT NUMBER: 166143
LDRD PROJECT TITLE: Time-Varying, Multi-Scale Adaptive System Reliability Analysis of Lifeline Infrastructure Networks
PROJECT TEAM MEMBERS: Jared Lee Gearhart and Nolan Scot Kurtz

ABSTRACT

The majority of current societal and economic needs world-wide are met by the existing networked, civil infrastructure. Because the cost of managing such infrastructure is high and increases with time, risk-informed decision making is essential for those with management responsibilities for these systems. To address such concerns, a methodology that accounts for new information, deterioration, component models, component importance, group importance, network reliability, hierarchical structure organization, and efficiency concerns has been developed. This methodology analyzes the use of new information through the lens of adaptive Importance Sampling for structural reliability problems. Deterioration, multi-scale bridge models, and time-variant component importance are investigated for a specific network. Furthermore, both bridge and pipeline networks are studied for group and component importance, as well as for hierarchical structures in the context of specific networks. Efficiency is the primary driver throughout this study. With this risk-informed approach, those responsible for management can address deteriorating infrastructure networks in an organized manner.

1. INTRODUCTION

Current US infrastructure needs are largely met by existing systems, e.g. power lines, transportation networks, water and gas pipelines, etc. However, the state of these systems is variable. Moreover, hazards such as seismic events, tornadoes, storm surge, and even corrosion, degrade the ability of these systems to perform. Due to the weakening of infrastructure and limited resources, governing bodies must determine which components are critical for the continuing support of network function. The American Society of Civil Engineers (ASCE) has rated the overall state of US infrastructure at D+, and the state of US bridges at C+ (ASCE 2013). An estimated $3.6 trillion (USD) would be required to improve this infrastructure to a level of "B" by 2020. While the grades for both general and bridge infrastructure have improved since the 2009 report card, the required investment has increased by $1.3 trillion, which is well above the inflation rate.

A related concern is that the average age of US bridges is 42 years. Because most bridges have a design life of 50 years, determining how structural response is affected by aging is particularly important. Since there are models for fragilities of bridges, pipeline sections, etc., one could view such components individually, but this would not accurately describe their role in the system. Since network analysis for these purposes rapidly becomes computationally expensive with the problem size, particularly for low probability events with simulation-based methods, e.g. crude-Monte Carlo Simulation (MCS), one must be careful with both network formulation and analysis. Additionally, there is a need to use field information in this process, either in constructing the initial state or updating the current state of the system, e.g. Bayesian parameter estimation or reliability updating. For these concerns, one must consider how to update the network state and how to perform the network analysis.

A framework that addresses these issues is developed, which accounts for deterioration, network complexity and size, component models, sources of correlations, efficiency issues, and adaptedness. After a major event, such as seismic hazards, infrastructure systems may become compromised in unintuitive scenarios. Because such networks are essential for emergency response, guaranteeing that the status of the network (i.e., the connectivity between special locations) is efficiently and accurately calculated is essential. A general framework relevant for performing this type of analysis, which is applicable to all infrastructure, is shown in Figure 1.

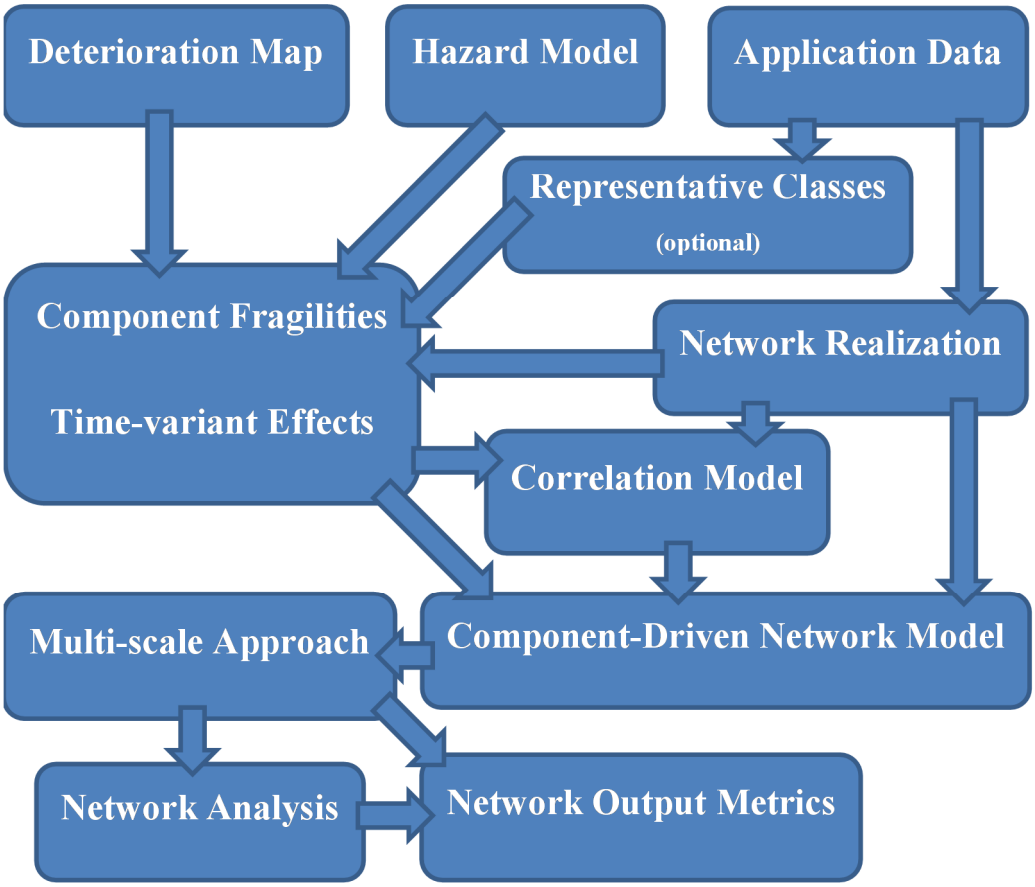

Figure 1. General Network Model Description.

There are three basic inputs to this framework. The first is the deterioration map, which determines how the deterioration parameters are spatially distributed. The second is the hazard model, i.e., the attenuation rule, the "hazard map." The last is the data specific to the application, e.g., bridge data, pipeline locations. In cases where the infrastructure contains a large range of components, e.g., bridge networks, representative classes may be used to represent groups of components to diminish computational expenses. The network realization describes the topology and how the distinct components (e.g. tunnels, culverts) are located within the topology.

Using these inputs, a component-driven network model is created using the following approach. While the component fragilities describe the likelihood of the limit states subject to hazard, e.g., seismicity and/or storm surge, the effects of time variance must also be handled. For this analysis, time variance is handled by "Fragility Increment Functions" (FIFs), which rescale the fragilities using ratios that account for the structural configuration and the environmental factors. The rest of the network model is driven by these high-information-content component models, resulting in a component-driven network model. Accounting for the individual components and the network realization, the sources of correlation between components failures are directly found. Using all of these inputs and techniques, a component-driven network model is created.

Multi-scale approaches (i.e., those with many components per link) using hierarchical structures to account for scale, are used together with the network analysis, (i.e., the selective-Recursive Decomposition Algorithm, or "S-RDA") to diminish computational costs. Selected outputs identify the important aspects of the network model. These outputs include component importance measures, e.g., Conditional Probability Importance Measures, and the network disconnection reliability. When a multi-scale, hierarchical approach is used, the outputs identify a hierarchy of component groups and associated group importance metrics (GIMs).

The next three sections of this report demonstrate the key features of this framework that were developed during this effort, through five analysis examples. The first example demonstrates the concept of making such a methodology adaptive as new information from the field becomes available. Second, an example of the methodology is shown that uses a multi-scale, time-variant, realistic, seismic, bridge network. In this example, representative bridges are used to define classes into which the bridges are grouped, with many representative bridges located on each link. In the third example, a multi-scale approach that uses hierarchical structures to diminish the computational costs of network analysis is demonstrated. The fourth example explores the further uses of such hierarchical structures for describing the network. In the final example, a case study using a time-variant pipeline is presented. The final two sections of this report contain the anticipated impacts of this research and its conclusions.

1.1 NOMENCLATURE

ASCE	American Society of Civil Engineers
AHSIA	Automated Hierarchical Structure Identification Algorithm
Caltrans	California Department of Transportation
CE	Kullback-Leibler Cross Entropy
CE-AIS-GM	Cross Entropy-based Adaptive Importance Sampling using Gaussian Mixture
CE-AIS-SG	Cross Entropy-based Adaptive Importance Sampling using Single Gaussian
c.o.v.	Coefficient of Variation
CPIM	Conditional Probability Importance Measures
FIF	Fragility Increment Function
FORM	First-Order Reliability Method
GIM	Group Importance Measure
IS	Importance Sampling
iHL-RF	improved Hasofer-Lind Rackwitz-Fissler
JFK	John F. Kennedy International Airport
LA	Los Angeles Metropolitan Area
LAX	Los Angeles International Airport
LP	Linear Programming
MCS	Monte Carlo Simulation
NA	Not Applicable
Ncut	Normalized cut
PAM	Partition Around Medoids
PGV	Peak Ground Velocity
PSA	Pseudo-Spectral Acceleration
RC	Reinforced Concrete
RDA	Recursive Decomposition Algorithm
Sa	Spectral Acceleration
SORM	Second-Order Reliability Method
S-RDA	Selective-Recursive Decomposition Algorithm
US	United States
USD	United States Dollars

2. DETAILED DESCRIPTION OF EXPERIMENT/METHOD

2.1 CE-BASED ADAPTIVE USING A GAUSSIAN MIXTURE

An adaptive importance sampling (IS) approach is developed (Kurtz and Song 2013) by expanding an existing adaptive IS approach (Song, et al., 2006) that finds a near-optimal IS density by minimizing Kullback-Leibler Cross Entropy (CE) via pre-sampling (Rubinstein and Kroese 2004). Here, CE quantifies the difference between the absolute best sampling density and the current IS density; however, the existing approach used a uni-modal distribution of statistically independent random variables. While this limits its use, such an approach is absent from structural reliability, although entropy maximization has been used. Such a CE-based approach allows the general ability to find an optimal IS density for both structural reliability, which is a more general field that seeks to obtain the probability of the system failure event of various engineering applications, and Bayesian inference (Box and Tiao 1992). For such situations, a new adaptive IS approach is developed by incorporating a nonparametric multimodal density function, i.e., a Gaussian mixture (Bishop 2006), into the aforementioned CE approach. The background for this new approach is discussed below.

Because management entities handle a large amount of new data over time, network methodologies must be employed to manage that data. Bayesian inference is used to update a network methodology. While there are many nuances with this approach, it uses two distributions of data: 1) the "prior" distribution, and 2) the distribution of experimental observations. Typically, importance sampling (IS) is used. IS is a variance reduction technique that presents a significant computational benefit to crude-MCS using an alternative density (Shinozuka 1983, Engelund and Rackwitz 1993, Melchers 1999). However, there may arise a situation in which the "prior" has non-negligible information content, and the distribution centers for the prior and the experiments are not near one another. For this case, the optimal IS density would lie at an ambiguous location between the densities. To handle such cases, the general adaptive IS method described earlier was developed for structural reliability problems. This method does not require application-specific information, and performs well with typical applications. Furthermore, this approach allows the optimal IS density for Bayesian updating to be found adaptively.

Typically, structural reliability problems are attempted using the First- or Second-Order Reliability Method (FORM or SORM) (Der Kiureghian 2005). First, the point of maximum likelihood in the failure domain, i.e., the "design point," is located, which is typically found using a nonlinear constrained optimization algorithm (e.g. the improved Hasofer-Lind Rackwitz-Fissler (iHL-RF) algorithm (Zhang and Der Kiureghian 1995). FORM or SORM is used to approximate the failure domain with a linear half-plane or paraboloid, respectively; however, these methods may not always function appropriately, indicating that sampling methods ought to

be used. Due to its straightforward application, the first attempted sampling method is crude-MCS; however, the computational costs for those rare events described by computationally expensive limit state functions, e.g., those relying on finite element analysis, may be exceedingly large. IS is then attempted. To find an IS density, a Gaussian density is used, for which the mean vector is located using iHL-RF (Fujita and Rackwitz 1988, Melchers 1989). However, this approach cannot address either multiple design points (Der Kiureghian and Dakessian 1998) or the numerical issues affecting such an optimization (Liu and Der Kiureghian 1991).

To this end, a novel adaptive IS approach has been developed. The aforementioned adaptive IS approach is used to 1) find the regions that contribute most significantly to the failure event, and 2) compute the failure probability while minimizing the variance. Section 2.1.1 presents the rules that specifically govern the proposed method.

2.1.1 METHODOLOGY

This approach extends the method developed by Rubinstien and Kroese (2004). (For a summary of this approach see Kurtz and Song, 2013.) The Kullback Liebler CE, $D(f(\mathbf{x}),h(\mathbf{x}))$ between two functions $f(\mathbf{x})$ and $h(\mathbf{x})$ is defined as

$$D(p^*(\mathbf{x}),h(\mathbf{x};\mathbf{v})) = \int p^*(\mathbf{x}) \ln p^*(\mathbf{x}) d\mathbf{x} - \int p^*(\mathbf{x}) \ln h(\mathbf{x};\mathbf{v}) d\mathbf{x} \tag{1.1}$$

where $p^*(\mathbf{x})$ is the optimal density, and h($\mathbf{x}$;$\mathbf{v}$) is the IS density with parameters $\mathbf{v}$. Due to the form $p^*(\mathbf{x})$, it cannot be sampled directly, but can be found through optimization. The resulting optimization one uses to obtain updating rules for $h(\mathbf{x};\mathbf{v})$ with i-th variable values, becomes

$$\frac{1}{N}\sum_{i=1}^{N} H(\mathbf{x}_i) W(\mathbf{x}_i;\mathbf{u},\mathbf{w}) \nabla_{\mathbf{v}} \ln h(\mathbf{x}_i;\mathbf{v}) = 0 \tag{1.2}$$

where $\mathbf{x}_i$, i=1,…,N are sample values; $H(\cdot)$ is typically the indicator function $I(\cdot)$, taking a value of 1 when the limit state function is not positive, $g(\mathbf{x}_i)\leq 0$, and 0 otherwise; and $W(\mathbf{x}_i;\mathbf{u},\mathbf{w})$ is the ratio between the nominal unimodal uncorrelated multivariate Gaussian distribution, with parameters **u,** and the adaptive IS density with parameters **w** from the previous updating step. Using the Gaussian mixture model for $h(\mathbf{x}_i;\mathbf{v})$, the IS density has the form

$$h(\mathbf{x};\mathbf{v}) = \sum_{k=1}^{K} \pi_k N(\mathbf{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \tag{1.3}$$

where $N(\cdot)$ is a multivariate, Gaussian kernel, and $\pi_k, \boldsymbol{\mu}_k$, and $\boldsymbol{\Sigma}_k$ represent the k-th Gaussian kernel's proportion in the Gaussian mixture, mean vector, and covariance matrix, respectively, which correspond to the distribution parameters, $\mathbf{v}$. Because the gradient in Eq. (1.2) is appended to the summation in Eq. (1.3), the concept of latent variables is used to obtain kernel parameter updating rules, leading to the updating rules presented in Kurtz and Song (2013).

The algorithm can be presented as follows:

1. Initialize: t=0. Choose initial values of parameters $\pi_k^{(t)}, \boldsymbol{\mu}_k^{(t)}$ and $\boldsymbol{\Sigma}_k^{(t)}$, k=1,…,K.
2. Pre-sample: t=t+1. Generate N random samples $\mathbf{x}_1,\ldots,\mathbf{x}_N$ using "ancestral" sampling and parameters from t=t-1. Calculate the ρ-quantile of $g(\mathbf{x})$. ρ corresponds to the proportion of $\mathbf{x}_i$ where $g(\mathbf{x}_i)\leq 0$.
3. Update: Obtain $\pi_k^{(t)}, \boldsymbol{\mu}_k^{(t)}$ and $\boldsymbol{\Sigma}_k^{(t)}$ from updating rules.
4. Check convergence: If ρ<0.10, return to step 2. Otherwise proceed to step 5.
5. Final importance sampling: Estimate the failure probability $\hat{I}_t$

$$\hat{I}_t = \frac{1}{N_f}\sum_{i=1}^{N_f} I_{\{g(\mathbf{x}_i)\leq 0\}} W(\mathbf{x}_i;\mathbf{u},\mathbf{v}^{(t)}) \tag{1.4}$$

where N_f corresponds to a target c.o.v., typically 5%.

An illustrative example for a series system with multiple component limit states follows in Section 3.1. Further examples demonstrating the breadth and depth of this approach can be found in Kurtz and Song (2013). In Section 3.1, the following limit state function, as presented originally in Waarts (2000), is used to demonstrate this approach. The limit state function is represented by

$$g(\mathbf{x}) = \min\begin{Bmatrix} 3+(x_1-x_2)^2/10-(x_1+x_2)/\sqrt{2} \\ 3+(x_1-x_2)^2/10-(x_1+x_2)/\sqrt{2} \\ x_1-x_2+7/\sqrt{2} \\ x_2-x_1+7/\sqrt{2} \end{Bmatrix} \tag{1.5}$$

where X_1 and X_2 represent two uncorrelated standard Gaussian random variables. Figure 2 shows the limit-state surface and contours of a function proportional to the $p^*(\mathbf{x})$ in the standard Gaussian uncorrelated space of random variables. Note that this density indicates four separate areas of importance. The results for this experiment were N=10^3 and K=4. The proposed method, termed "AIS-CE-GM," was compared to crude-MCS, termed MCS, and the proposed method using a single Gaussian kernel, termed "AIS-CE-SG."

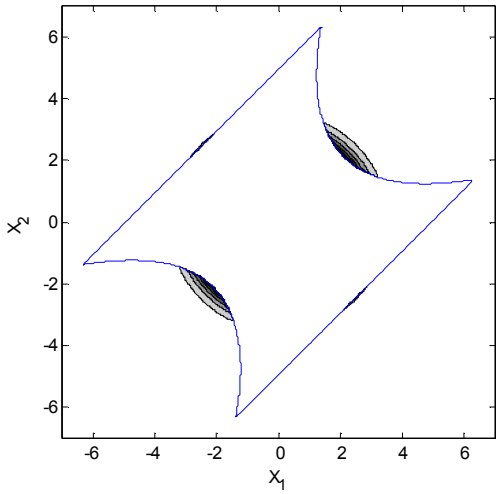

Figure 2. Optimal Density Shape.

2.2. SEISMIC RELIABILITY ANALYSIS OF DETERIORATING REPRESENTATIVE U.S. WEST COAST BRIDGE TRANSPORTATION NETWORKS

The methodology in Figure 1 has been applied to a deteriorating bridge network subject to seismicity (Kurtz, et al.; under review). This approach uses application data from Caltrans (California Department of Transportation) to locate highway bridges and to describe highway bridge structures. A clustering technique is used to find representative bridge classes specific to the network of choice. Bridge correlations are developed accounting for spatial seismic intensity, structural configuration, construction date, and deterioration. A multi-scale approach is used for each link to represent many bridges and to simplify the correlation calculation. The time-variant network disconnection probabilities and component importance measures are found using S-RDA.

To diminish sensitivity to rare event probabilities, Lim and Song (2012) proposed the S-RDA, which improves the original RDA by identifying critical sets first. Nevertheless, a multi-scale analysis is still required due to network size limitations. If such analyses need to be repeated with time-variant deterioration and retrofit data, computational costs increase. Several analysts have attempted to define such deterioration by analyzing specific future time points (Liu and Fragopol 2005, Guikema and Gardoni 2009, Lee, et al., 2010). Many analyses used simplifications, e.g., HAZUS fragilities, simplistic capacity degradation, etc. A primary criticism for such approaches is that they are supported by engineering judgment, without accounting for data.

To address concerns associated with the factors discussed above, probabilistic seismic capacity and demand models for both single-bent and multiple-bent reinforced concrete (RC) bridges have been developed using experimental data and time-history finite element analysis as input to a Bayesian estimation procedure (Gardoni, et al., 2002; Gardoni, et al., 2003). Furthermore, time

variant deterioration functions that reshape as-built fragilities while accounting for structural configuration and atmospheric condition have also been developed (Gardoni and Rosowsky 2009). The exact methodology of the proposed bridge network approach is presented in Section 2.2.1, below.

2.2.1 METHODOLOGY

To guarantee that accurate bridge models are used without incurring too much computational complexity, representative bridge classes will be used that have high information-content, time-variant bridge fragility models available. These models are based on Partition About Medoids (PAM) clustering (Kaufman and Rousseeuw 1990) of specific data for an area of interest, e.g., the Los Angeles Metropolitan Area (LA) (CalTrans 2013). This data was appropriately filtered and compiled. This novel approach is more specific than arbitrary categories would allow. The optimal number of clusters for this analysis (seven) was determined using the gap heuristic (Tibshirani, et al., 2000). Furthermore, the resulting clusters were filtered to three, which accounted for 70% of the total bridges in the dataset. Of these three classes, the first was a two-span bridge, while the last two were three-span bridges. Because the probabilistic capacity model was section-based (Choe, et al., 2007), it could be used for all bridges classes. The probabilistic demand model for the two-span bridge, which only has one bent, lent itself to closed form fragilities (Huang, et al., 2010), as shown in Figure 3. Note that this bridge class fragility is bi-variate, requiring two seismic intensity inputs: normalized Peak Ground Velocity (PGV) and normalized Pseudo-Spectral Acceleration (PSA). However, each of the other classes included a series system of bents, requiring crude-MCS sampling of the structural design variables and Bayesian fragility coefficients (Gardoni, et al., 2003), as shown in Figure 4. Note that these bridge class fragilities require a single input of Spectral Acceleration (Sa). (For a full description of seismic intensity measures, review Elnashai and Di Sarno (2008).) The as-built fragilities were then reshaped over time using “Fragility Increment Functions” (FIFs), which are ratios that model both the increased variance and fragility with further chloride-induced RC deterioration. FIFs are also tailored to the bridge’s structural configuration and environment.

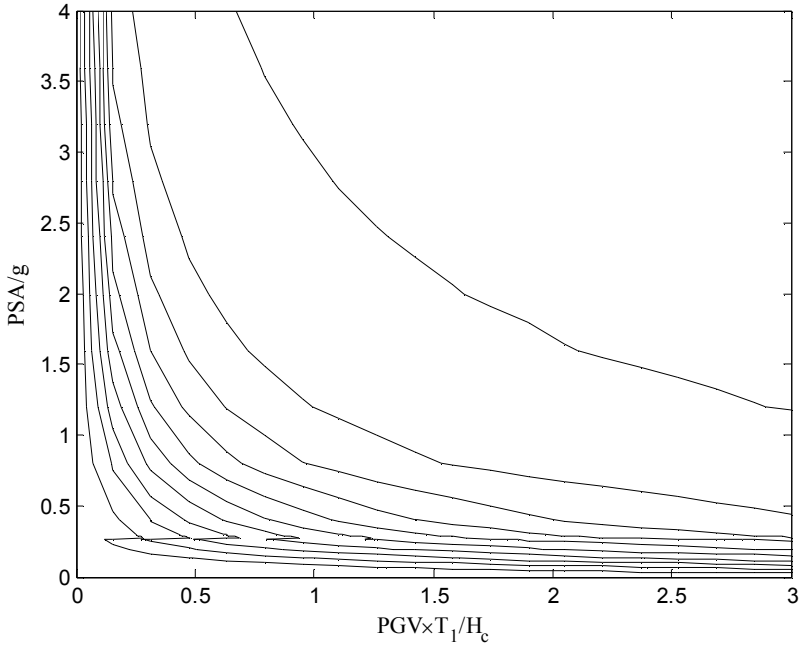

Figure 3. Bridge Class 1 Fragility.

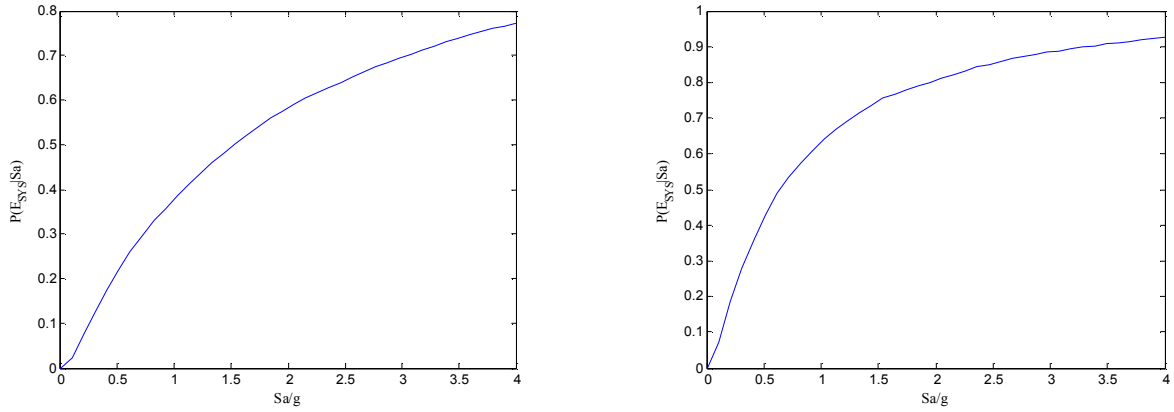

Figure 4. Fragilities for Bridge Classes 2 (left) and 3 (right).

Because the bridge models account for only the reliability of such components, the correlations, multi-scale representations, and parsing of seismic intensity parameters are described. These were developed following the suggestion of Lim and Song (2012) in a study for pipeline networks, but were specially developed for both bridges. The two sources of ground motion correlation come from the inter-event residual and intra-event residual, which account for spatial correlation in a seismic event; however, due to the method used to obtain the bridge failure correlations, structural configuration, construction date, and deterioration are also accounted for. Since S-RDA is strongly sensitive to size and the end goal of this study is to analyze thousands

of bridges, links containing several bridges representing major state highways are used, requiring some basic usage of set theory. To obtain link reliabilities as series systems of bridge reliabilities, an efficient, high-dimensional, multivariate normal, numerical integration scheme is required (Genz 1992). Furthermore, to obtain link failure correlations, the process used to obtain bridge failure correlations from ground motion correlation still applies with slight modification; nevertheless, when many links contain hundreds of bridges, several thousand bridge failure correlation terms are relevant to each link failure correlation term. To make this process more efficient, only intra-link bridge failure correlation terms are calculated with the full process, while the inter-link bridge failure correlation terms are calculated using an interpolation between three distributed locations on each link. Given the component description, the network analysis and output metrics are discussed.

To summarize the network analysis algorithm and its typical output, the S-RDA is a bounds-based convergence algorithm for network disconnection events. S- RDA improves upon the RDA by using a version of Dijksta's algorithm that maximizes the product of the reliabilities. This obtains the maximum likelihood disjoint link-set, which represents a safe path from the source to the terminal node, updating the upper bound on the system reliability using simple subtraction. After decomposing this identified set, further link-sets, or cut-sets, which disconnect the source and terminal nodes, are identified by the degree to which they influence the network disconnection until a desired tolerance is found, e.g., 1% or 5%. Using these identified disjoint sets, the link Conditional Probability Importance Metrics (CPIMs) can be calculated using combinations of summations. CPIMs are formulated as the conditional probability that a component has failed given the network has failed, which shows the extent to which the component of interest participates in the network disconnection event. The CPIMs have both cut-set and link-set representations that use a subset of identified sets for converged results with no need for further reliability analysis. It must also be noted that the network represents both nodes and links as node-type components. The original nodes in this network are modeled as infinitely reliable and as uncorrelated with the original links. Given this network analysis approach, a network of choice is investigated.

An example based on the LA's network complexity is explored in Section 3.2. (See Figure 5 for a description of the network to be analyzed.) As opposed to locating each bridge in the network individually, which is a highly time intensive process, the number of bridges on a link is found by link and global averages from the Caltrans data, and the class of each bridge in a link is determined based on a multinomial distribution sampling of the bridge classes and their proportion of the Caltrans data. The epicenter of the Northridge 1994 earthquake, represented with a star, is used for the seismicity, with an increased magnitude of 8.0. An attenuation rule is used in conjunction with this (Boore and Atkinson 2008). The double-dash dotted lines represent subjunctive links, which are infinitely reliable, uncorrelated, and are used to model many source

and terminal nodes. Specifically, these are located on evacuation paths, which are connected to the terminal node,

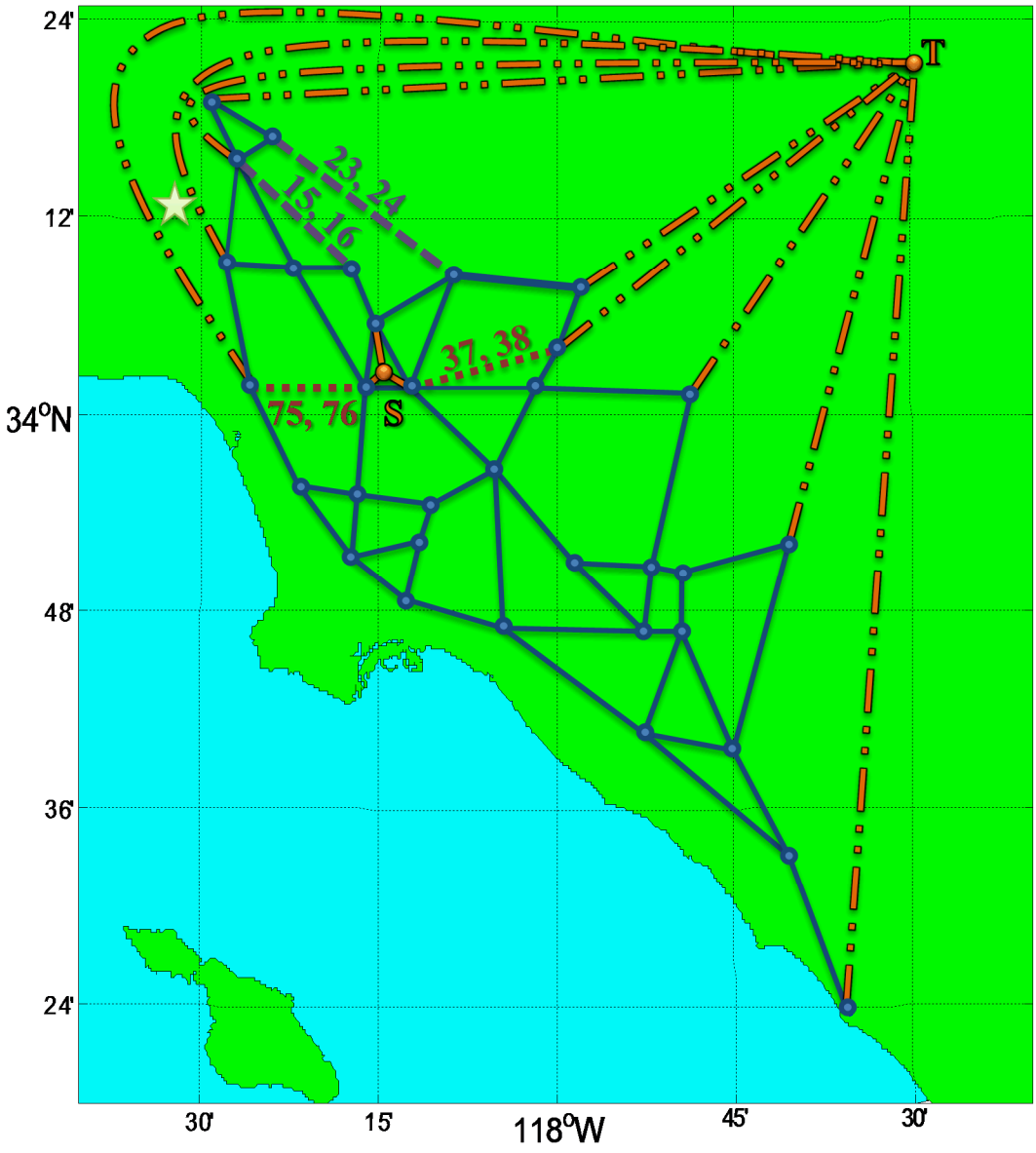

Figure 5. LA Network Representation.

represented with T, as well as on downtown paths, which are connected to the source node, represented with S. The scale of the network is shown with latitude and longitude coordinates. Several deterioration scenarios are compared and contrasted using crude-MCS: the as-built state, where no deterioration has occurred, the 2013 case, where all bridges are deteriorated to 2013 from their build date, the 100-year Deteriorated case, where all bridges have been deteriorated 100 years from when the last class was constructed, and the 100-year Retrofitted case, where the

same conditions for the 100-year Deteriorated case (except the first bridge class has been retrofitted to as-built conditions). To investigate the sensitivity to the multinomial distribution sampling, each deterioration scenario uses a different bridge location. The CPIMs were obtained such that median difference between link-set and cut-set definitions was 2%. The links corresponding to the top ranked link CPIMs are also shown on Figure 5, as will be explained in Section 3.2.

2.3 MULTI-SCALE SEISMIC RELIABILITY ANALYSIS OF LARGE INFRASTRUCTURE NETWORKS USING HIERARCHICAL STRUCTURES

As seen earlier in this report, size limitations must be addressed. The earlier bridge network analysis considered some efficiency improvements for the component model by developing a multi-scale approach for link reliability analysis of many bridges; however, the S-RDA must still be improved for networks having larger numbers of links. To this end, several efforts have approached multi-scale analysis in the form of a "divide and conquer" method, where subnetworks are extracted in a hierarchy and analyzed separately, starting at the lowest level. As the analysis progresses upward, the analyzed subnetworks are replaced with "super components." Der Kiureghian and Song (2008), attempted this using both visual inspection and the Linear Programming (LP) bounds method.

To remove the contingencies that come with defining the subnetworks, Gomez, et al. (2013), used a hierarchical paradigm in which clusters were replaced with infinitely reliable "super nodes". While this was helpful, the system disconnection reliability was heavily influenced, and correlations could not be accurately modelled. Lim, et al. (in print), proposed a method using a hierarchical clustering, in which subnetwork adjacent nodes were connected with "super-links" that modelled the subnetwork more accurately, and allowed correlations to be accurately modelled using Ncut-spectral clustering (Von Luxburg 2007) with S-RDA. Nevertheless, both of these approaches relied on limited heuristic approaches to specify the optimal number of clusters per analyzed subnetwork, and have not been tested on networks of large size. Clearly, automation may solve these issues; however, it must not propagate significant error for large hierarchies. Furthermore, automation must be developed with objective rules and must be tested on several networks.

Several such automated hierarchical structure identification algorithms (AHSIAs) have been developed and tested in the manner described above to address this (Kurtz, under review). A brief description of the five developed AHSIAs is provided in the section below. Lastly, a test example is described.

2.3.1 METHODOLOGY

Five AHSIAs were developed. The key features of these algorithms are summarized in Table 1. When developing an AHSIA, there are two cases that must be avoided: (1) super-link representation if it presents little difference from the original subnetwork, and (2) super-link representation if it requires too large a subnetwork be analyzed. Additional clustering can be used to handle the second case. Each AHSIA is defined in the paragraphs below. The convergence approaches describe whether the algorithm requires a global number of bottom-level clusters be identified (Global), or whether the algorithm develops each branch until a set of convergence criteria are met (Branchwise). "Force clustering" is the algorithmic requirement to divide a cluster when the smallest cluster representation is the original. The input parameters to AHSIAs are those inputs the user must specify beforehand, other than the network topology to run the algorithm. Lastly, the depth of each AHSIA describes how many levels of subclusters can be typically expected.

Table 1. AHSIA Qualitative Description

Algorithm		Convergence	Forced clustering	Parameters	Depth
I-AHSIA	Bi	Branchwise	Yes	1	Deep
	min	Branchwise	Yes	1	Deep
RI-AHSIA		Branchwise	No	2	Moderate
N-AHSIA	Stage	Global	Yes	1	Shallow
	Global	Global	Yes	1	Shallow

The first AHSIA was termed "Ignorant-AHSIA" or I-AHSIA. This name was justified, because it only analyzed the subnetwork structure and did not take further information from network objective functions (although Ncut-spectral clustering is still being used). The convergence of I-AHSIA is branchwise, i.e., each branch must meet the convergence criteria before it is finished. This algorithm operates by finding the first division k of the subnetwork starting with k=2, such that the representation is less than some threshold *num* (e.g., 30), and the next largest division (k+1) is greater than the threshold. This optimization is posed this way so that the subnetwork analyzed by S-RDA is neither too large nor too small; however, some topologies do not allow the represented subnetwork to be less than *num*. Therefore, two heuristic relaxations are used when this optimization is not possible: (1) bifurcation, i.e., k=2, and (2) relative minimum, i.e., the minimum subnetwork representation. Convergence occurs when the subnetwork without super-links is less than or equal to *num*. Nevertheless, this algorithm will always select a division until all "leaf" clusters satisfy the convergence.

Because this methodology clearly presents some issues for the I-AHSIAs, a "Relative Ignorant-AHSIA" or RI-AHSIA is proposed. This optimization is purely relative and selects the minimum

representation. Additionally, the optimization considers k=1, or the size of the original subnetwork without super-links. This means that forced division no longer occurs. Additionally, the difference between the original and the represented subnetwork is guaranteed to be above some threshold, e.g., 10. This is relaxed for k=2,3, so that a larger diversity is considered. If such relaxations of the difference constraint occur within 3 levels of the leaf subnetworks, the tree will be "pruned," because this relaxation provides little benefit. All five AHSIA use such pruning. There are two convergence criteria: (1) a non-reducible branch, i.e., k=1, and (2) a reducible branch, i.e., the original subnetwork is less than some threshold, such as 30. Note that the first three AHSIA are "ignorant" of specific network objective function information, other than Ncut-spectral clustering.

To handle the network objective function information explicitly, "Normalized cut (Ncut)-based-AHSIA," or N-AHSIA, is used. N-AHSIA uses global convergence, as opposed to the branchwise convergence of the three earlier examples. This algorithm continues until $k^{\max}$, the number specified by minimum value of the modularity objective function (Laarhoven and Marchiori 2013), is reached. At each level of analysis, certain further divisions, with k corresponding to the minimum Ncut value, are considered. Either all divisions in the same stage, "stagewise," or all divisions globally, "global," are considered. Based on whether the stagewise or global versions are used, N-AHSIA selects the certain, further division with the largest difference between the represented and the original subnetworks.

Each of these clustering techniques results in a network "dendogram" (Hastie, et al., 2009), which corresponds to a "tree" description of the hierarchical structure for the network. Especially for networks which have clear hierarchical structures, the form of this dendogram is essential for evaluating how these AHSIA perform. Examples of a RI-AHSIA dendograms are shown in Figures 11, 12, and 15. The numbers at the leaf clusters, which are at nodes with only one connection, and the numerators at other locations correspond to the size of the original subnetwork, which is the number of relevant nodes and uni-directional links. The denominator adds the super-link represented cluster sizes with the adjacent uni-directional links to obtain the size of the represented subnetworks. The super-link represented cluster sizes are the sum of adjacent nodes, source and/or terminal nodes, and uni-directional super-links.

For a case study, a possible realization of US pipeline network is investigated in the Section 3.3, as shown in Figure 6. The pipeline component model in Lim, et al. (in print), is used for this example. The Los Angeles International Airport (LAX) and the John F. Kennedy International Airport (JFK) are selected for the source and terminal nodes, denoted by 9 and 67. All nodes used in this network correspond to airports with at least 10 uni-directional connections (Opsahl 2011). A "nearest neighbor" graph is constructed by connecting a given node to two of the six closest nodes based on uniform random sampling. This network has 103 nodes and 183 bi-directional links for a total of 469 components. Several numbered links are also shown in Figure

6, which correspond to the top-ranked CPIMs. These will be discussed in Section 3.3. Note that several scaled parameters must be specified for the multi-scale S-RDA. For the bounds tolerance, the following equation is used:

$$tol_{bounds} = \begin{cases} 0.01, & UB > 0.10 \\ 0.01\dfrac{UB}{0.10}, & 0.10 \geq UB > 0.005 \\ 0.005, & 0.005 \geq UB \end{cases} \tag{3.1}$$

where UB represents the current value of the upper bound for the current analyzed subnetwork. For calculating the correlation terms, the following equation is used:

$$tol_{corr} = \begin{cases} 0.30, & UB > 0.10 \\ 0.30\dfrac{UB}{0.10}, & 0.10 \geq UB > 0.02 \\ 0.02, & 0.02 \geq UB \end{cases} \tag{3.2}$$

Furthermore, the seismicity is modelled based on the New Madrid seismic zone, with a moment magnitude of 8.0, and denoted by a star, which corresponds to the relevant earthquakes in the 1800s. The most important CPIM components, as well the dendograms, are considered, in addition to the comparison between AHSIAs.

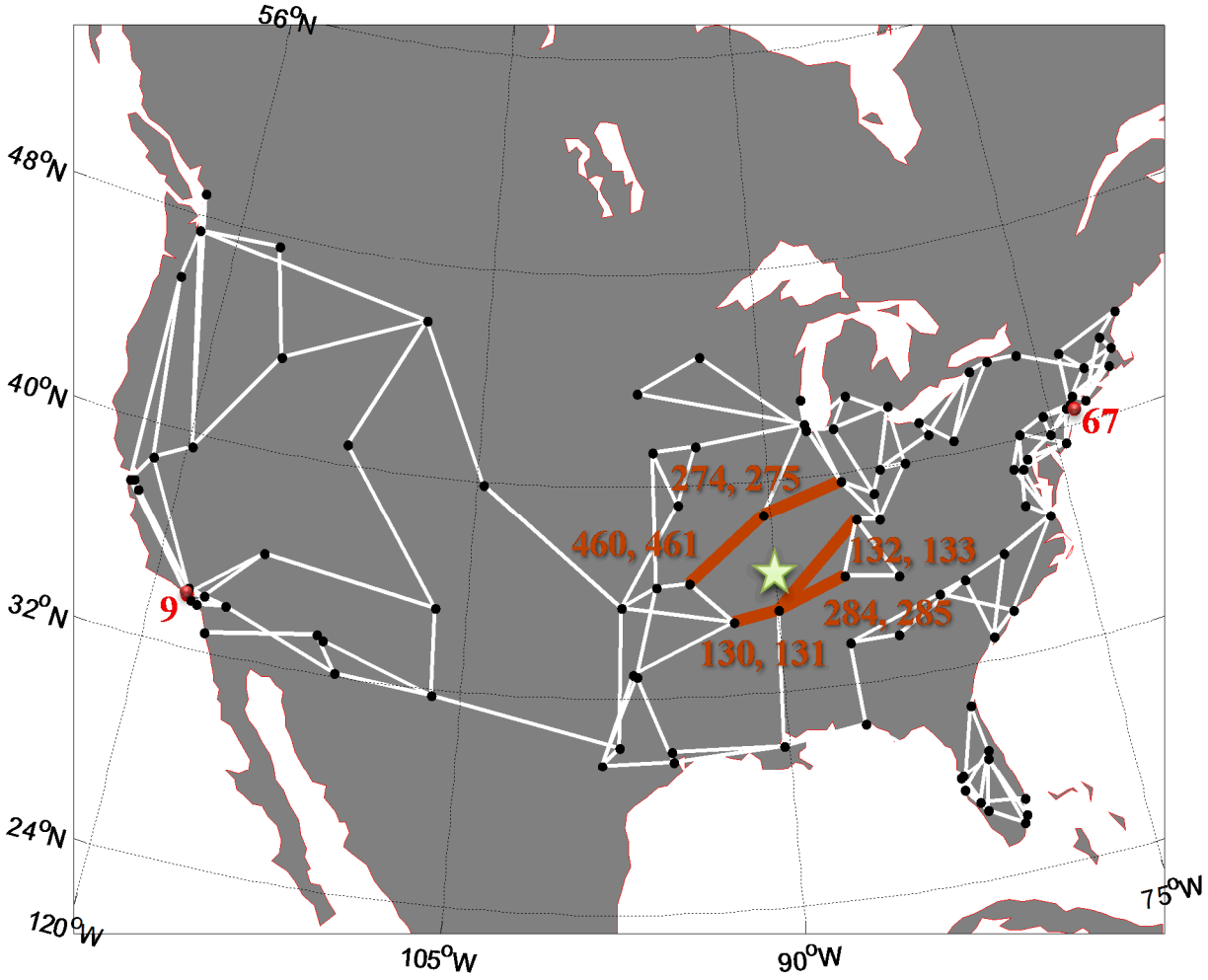

Figure 6. US pipeline network representation.

2.4 GROUP IMPORTANCE AND REGIONAL MANAGEMENT STRATEGIES (GIMS)

Group importance measures (GIMs), using union and intersection descriptions with link-set and cut-set information from the network analysis, are developed in this section. These methods have been investigated in depth in Kurtz (under review), and analysis based on these investigations will be applied to a case study, as described in Section 2.4.1. These regionally based approaches are not affected by the same issues that influence component-level importance metrics.

As seen in earlier studies, there are many reasons to perform network studies. To improve networks, defining the vulnerabilities with reference to the overall network function is essential; however, the vast majority of work focuses on the component level. Component views may be very sensitive to correlations that elevate the importance of components, but have little effect on

network function. Furthermore, these views may also be very sensitive to specific terminal and source node combinations, as well as to specific seismicity. Furthermore, to manage such networks, it may be more useful to specify regional strategies based on groupings. As seen in Section 2.2 of this report, the effects of deterioration were rather erratic at the component level. GIMs will be investigated as a way to manage these concerns.

A desire to consider more than one component at a time in terms of rankings is something seen in many fields. In finance, many described risk to portfolios using "price risk" in the 1990s. This construct is defined as "the likelihood of value loss for an institution's entire portfolio based on fluctuations in market conditions, e.g., commodity pricing, interest rates." For traditional products, such price risk is separable, meaning that part of the magnitude for the portfolio risk from one of the risk factors can be found by performing sensitivity analysis for that risk factor. This assumes that such factors are independent; however, there exists the so-called "correlation product" where many risk factors are found to be non-separable in nature due to strong correlations. For these reasons, financial risk analysis must consider combinations of risk factors when analyzing correlation products (Mahoney 1995).

Such correlated risks indicate the structure in which the analysis should be conducted. In data mining, many have found that association rules are strongly useful. These are applications where the percentage of transactions containing two distinct products, e.g., diapers and beer, is mined from a data set (Holt and Chung 1999). Information about products with a high incidence of sale and strong association rules help to indicate better store layouts for vendors. In a similar manner, grouping components in a network help indicate the network structure. Now that such AHSIA are available to describe network groupings on several levels, an exact methodology for GIMs can be described.

2.4.1 METHODOLOGY

Taking inspiration from the CPIM definition, one can use "network function" to describe GIMs. Network function describes how flow can happen through a component, given the system failure. Extrapolating this to clusters, GIMs describe whether or not a path through the cluster is available. Using Boolean descriptions, there are two extremes one can use to describe the identified groups: intersection, or union. The intersection description is a joint failure and parallel system. It conservatively describes the cluster failure; therefore, one would expect that it favors smaller values and smaller groupings of components. It is defined as "the conditional probability that all relevant nodes and links fail given the network has become disconnected".

The union description only requires that at least one relevant node or link fail. This description is also a series system. Because it is a more relaxed description of the cluster failure, one would expect that it favors larger values and larger groups of components. The union description is

defined as “the conditional probability that at least one relevant component fails given the system failure”. Both descriptions have link-set and cut-set definitions, as shown in Kurtz (under review); however, convergence is typically not an issue due to using the highest level super-link representation.

These GIMs are evaluated using a refined LA traffic network representation in Section 3.4, as shown in Figure 7. This network has 77 nodes and 117 bi-directional links for 311 total components. The seismic event, denoted by a star, is modeled after the Northridge 1994 earthquake, and assumes a moment magnitude of 6.7. The source and terminal nodes, placed at 50 and 76, are chosen to be at the largest extent of the network. The parameter in Equation 3.2 had to be changed from 0.30 to 0.10, because the first value resulted in a convergence that was too coarse. Because RI-AHSIA performed best in Kurtz (under review), it was used for this example. The component CPIMs, the top 10 of which are shown in Figure 7, will also be investigated and compared.

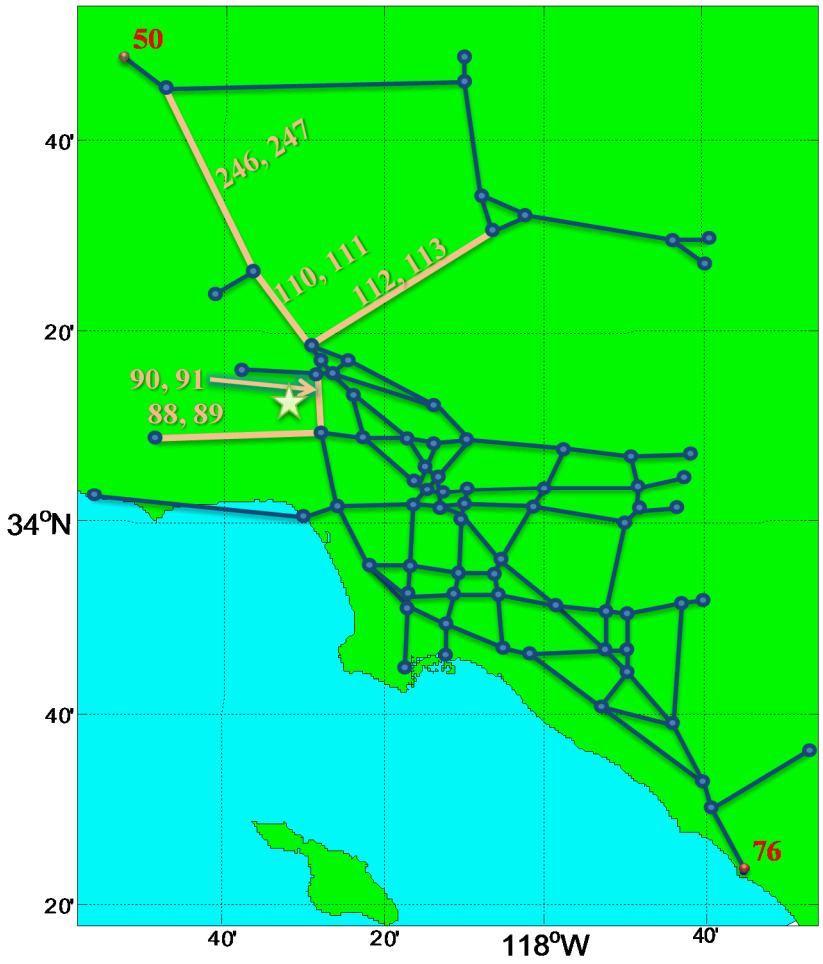

Figure 7. Refined bridge network representation.

2.5 DETERIORATING EFFECTS ON PIPELINE NETWORKS

To investigate a different network subject to both deterioration and time-variant effects, a deteriorating pipeline model subject to non-uniform deterioration is presented. Using a fitted FIF, the pipeline fragilities are deteriorated over time. The deterioration map is created by having the components closest to the seashore begin deteriorating earliest, while the ones furthest away begin to deteriorate latest. Three time points are considered and compared. Additionally, a reliability-based hierarchical structure, which may become time-variant depending on changes to link reliabilities, is compared to a connectivity-based one. The changes over time for component importance are also investigated.

An important attribute of existing civil infrastructure is that it is consistently deteriorating over the life-cycle, making the state of repair and the need for investment in infrastructure inherently dynamic. In the broader field of deteriorating bridge network seismic risk analysis, some have attempted analysis using data-driven statistical learning concepts (Rokneddin, et al., 2012); however, no one has attempted this type of analysis using the analytical methods used in the development of this report. Unfortunately, such statistical learning approaches tend to neglect underlying physics and analytics, while the approaches used in this report model these explicitly.

The FIFs may also be generalized to reshape the brittle pipeline. A study for investigating non-uniform deterioration using the S-RDA disconnection bounds, component CPIMs, and reliability-based hierarchal studies allows one to consider how much deterioration affects network characteristics.

2.5.1 METHODOLOGY

Using the same pipeline network model as in Section 2.3, the major differences that will be discussed are the deterioration model and the distribution of deterioration parameters. Using the same approach for the bridge model, FIFs relevant to the pipeline model are obtained. Because appropriate data is not available, the original forms in Gardoni and Rosowsky (2009) are considered. The general FIF form is

$$FIF_{\delta v}^{pipe}\left(t,D\right)=\frac{1+\left(a/D\right)^{b}}{1+\left\{\left[a-c\left(t-T_{corr}\right)^{d}\right]/D\right\}^{b}},\qquad t\geq T_{corr} \tag{5.1}$$

where D is a seismic intensity parameter, a, b, c, and d are variables that control the shape, and t is the time since the start of analysis. The shape parameters are tuned such that the steepest analyzed fragility, i.e., the node fragilities in the pipeline models, have less than a 1% survival area at t=50, chosen because most design lives are 50 years. To ensure the survival area is small enough, the time-variant normalized average, $\overline{A_{frag}}(t)$, was used. This is determined by

$$\overline{A_{frag}}(t) = \frac{1}{d_{\max}} \int_0^{d_{\max}} FIF_{\delta v}^{pipe}(t, D) F_{\delta v}^{pipe}(D) dD \tag{5.2}$$

where $F_{\delta v}^{pipe}(D)$ are the node fragilities. To select the shape parameters, $1 - \overline{A_{frag}}(t) < 1\%$ satisfies the earlier condition. Through this fitting, a=1.9, b=4.0, c=0.032, d=1.1, and T_{corr}=10. Revisiting the non-uniform deterioration map, the readily modifiable terms relevant to deterioration only impact the start of corrosion, not the rate. For those reasons, corrosion initiation times, T_{corr}, are exponentially distributed from 1 "year" for the 10% of components closest to the shore, to 20 "years" for the 10% of components furthest from the shore, where the 1 "year" and 20 "years" corrosion time correspond to splash-zone and atmospheric environmental characteristics, as shown in Figure 8.

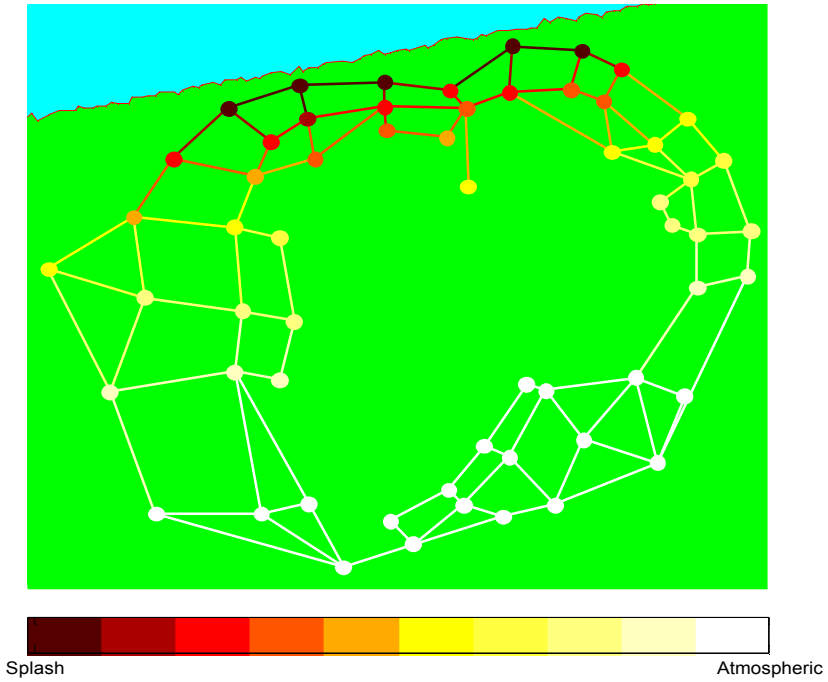

Figure 8. Pipeline deterioration classification.

Lastly, to introduce time-variant sensitivity into the hierarchical structure, *connectivity information*, which is used to generate the connectivity-based hierarchical structure, is replaced with *reliabilities* in the node adjacency matrix. In simpler terms, the matrix entries having a value of one are replaced with the relevant link reliability. Using such a matrix modification, the reliability-based hierarchical structure is obtained. The reliabilities decrease over time as the deterioration occurs, possibly influencing how the structural organization happens. This may present benefits for specific applications.

As a case study of deteriorating pipeline networks, an example from Lim, et al. (in print), is investigated in Section 3.5, as shown in Figure 9. This network has 59 nodes and 99-bi-directional links for 257 components. The seismic event, denoted by an epicenter at the star, has a moment magnitude of 6.0. The source and terminal nodes are placed at 16 and 46. Three time points are considered: t=0, 13, and 25 years. The top 10 ranked components are also shown in this figure, as will be discussed in the relevant results section. Furthermore, the network rubrics mentioned earlier, specifically, the variations between time points, will be used to measure the effects of deterioration using both connectivity and reliability-based hierarchical structures.

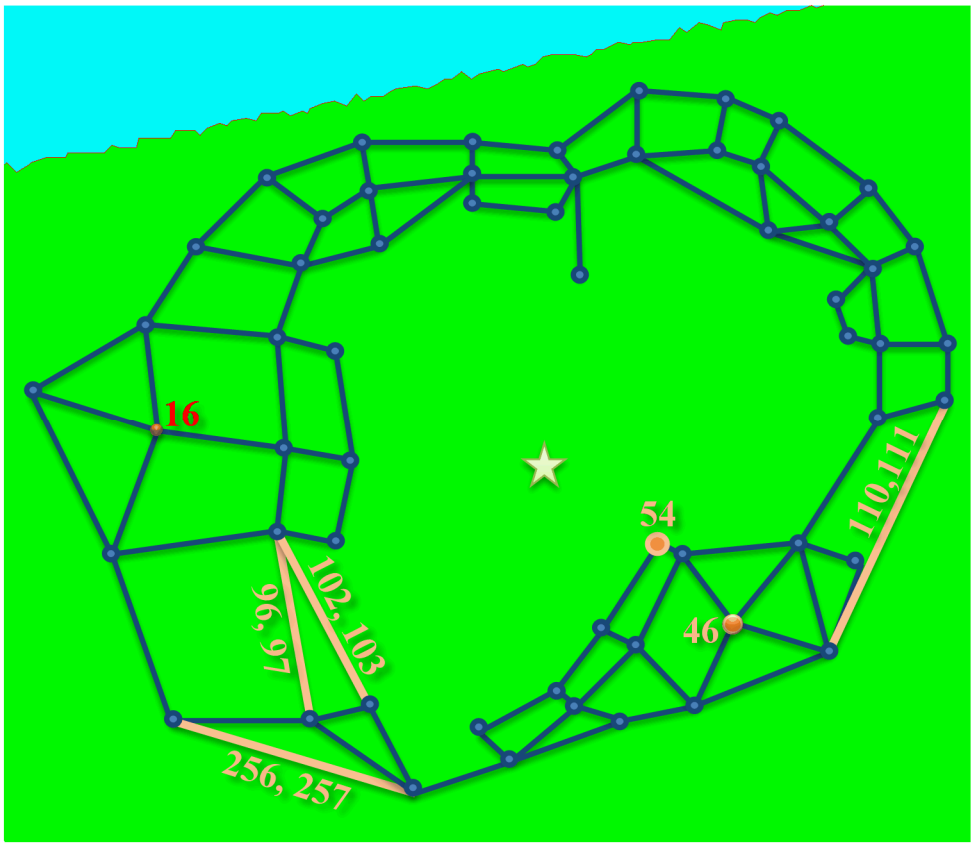

Figure 9. Pipeline network configuration with 0 year top raked components.

3. RESULTS

3.1 CE-BASED ADAPTIVE IS USING GAUSSIAN MIXTURE

Using the suggested procedure mentioned earlier, Table 2 shows the summary of results for several coefficient of variation (c.o.v.) values. An addition symbol is used in both the CE-AIS-SG and CE-AIS-GM columns to show the number of pre-samples added to the number of final IS samples. The single Gaussian approach required an extra step to converge in the search when compared to the Gaussian mixture approach. For all three c.o.v. values, the single Gaussian approach required an extra order of magnitude to converge for the final IS when compared to the Gaussian mixture approach.

Table 2. Comparison for Different Sampling Methods

c.o.v. (%)	Number of Samples			Failure Probability		
	MCS	CE-AIS-SG	CE-AIS-GM	MCS	CE-AIS-SG	CE-AIS-GM
10	60,000	4,000+500	3,000+30	1.83×10^{-3}	7.97×10^{-4}	1.50×10^{-3}
5	1.90×10^{5}	4,000+1,500	3,000+348	2.12×10^{-3}	8.80×10^{-4}	2.12×10^{-3}
3	5.20×10^{5}	4,000+2,500	3,000+943	2.16×10^{-3}	8.72×10^{-4}	2.15×10^{-3}

Figure 10 shows the convergence of the IS density, requiring 3 updating steps. No preliminary component reliability analysis is required, as is typical for existing series system approaches. Due to the non-circular shapes present, the effect of Σ_k at each region of importance can be seen. These shapes also adhere to the slopes of the component limit states. Note that the proposed method converges two orders of magnitude more quickly than crude-MCS, and that the single Gaussian approach exhibits false convergence. The latter point occurs due to the single Gaussian only being able to encapsulate one region. The c.o.v. behavior for this approach also exhibits jagged behavior with further samples, which would not manifest at larger target c.o.v. values.

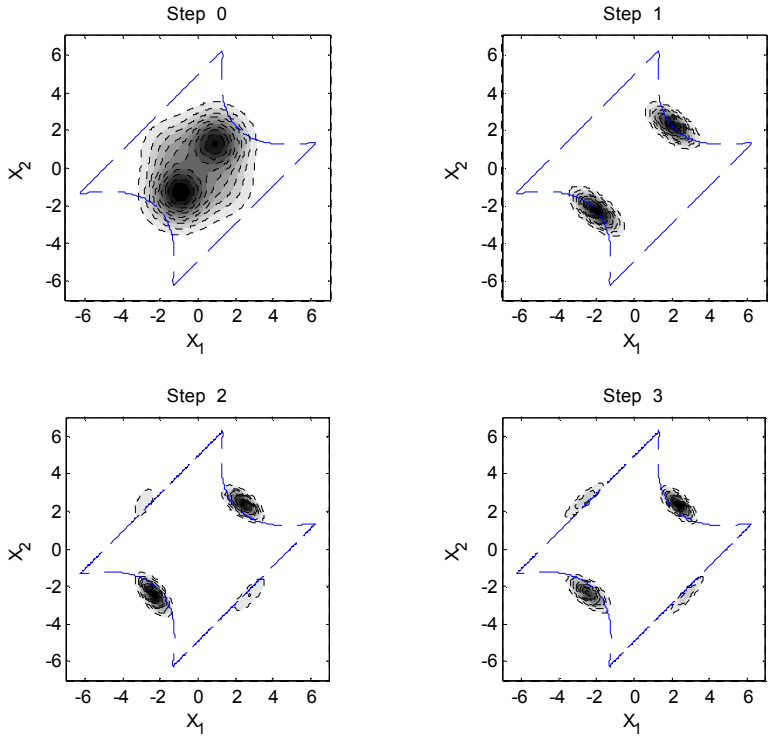

Figure 10. Convergence of Gaussian mixture.

3.2 SEISMIC RELIABILITY ANALYSIS OF DETERIORATING REPRESENTATIVE US WEST COAST BRIDGE TRANSPORTATION NETWORKS

Using the earlier parameters, the S-RDA bounds and crude-MCS estimate, P_f, with 5% c.o.v., were obtained in Table 3. The bounds represent the closest S-RDA bounds contained by the percentage described. Because the crude-MCS estimate is contained at 1% bounds, the best agreement occurs for the As-built and 100-year Deteriorated cases. In contrast, the 2013 case contains the sampling estimate at 2%, which is initially below the bounds, while the retrofitted case contains the sampling estimates at 3%, which is initially above the bounds. Due to the nature of the subjunctive representation, many critical link-sets are identified quickly, indicating that the upper bound converges quickly. This indicates that the 100-year Retrofitted case is particularly problematic. Furthermore, if cases using different bridge locations were investigated at 20-year intervals, non-monotonic increases in bounds values occur, indicating sensitivity to bridge locations.

Table 3. Crude-MCS Comparison: Disconnection Probability

Proposed Methodology					Crude-MCS
Case		10% bounds	5% bounds	1% bounds	P_f
As-built	UB	0.116	0.0912	0.0734	0.0721
	LB	0.0231	0.0415	0.0634	
2013	UB	0.140	0.114	0.0973	0.0840
	LB	0.0435	0.0647	0.0873	
100-year Deteriorated	UB	0.161	0.139	0.123	0.122
	LB	0.0639	.0891	0.113	
100-year Retrofitted	UB	0.143	0.125	0.106	0.113
	LB	0.0473	0.0747	0.0965	

The top 5 CPIM values that were obtained are shown in Table 4, with the execution times shown below the case names. The cut-set and link-set values in this table can be interpreted as how well the results have converged based off of the difference between them. Because the link-set values are probabilities greater than one, there is some residual error that coming from an efficiency approximation; however, this error is not excessive based on further analysis.

Table 4. Top 5 CPIMs for Validation Cases

Proposed methodology			
Case	Link	Cut-set	Link-set
As-built (29,600 sec)	76	0.997	1.07
	37	0.993	1.07
	75	0.995	1.06
	38	0.995	1.06
	23	0.98	1.04
2013 (22,014 sec)	38	0.995	1.06
	76	0.994	1.06
	75	0.995	1.06
	37	0.991	1.05
	15	0.981	1.04
Deteriorated 100 years (18,479 sec)	38	0.989	1.04
	37	0.988	1.04
	75	0.988	1.04
	76	0.988	1.04
	16	0.966	1.01
Retrofitted 100 years (44,442 sec)	38	0.997	1.05
	76	0.998	1.05
	75	0.996	1.05
	37	0.997	1.04
	24	0.987	1.03

While each case uses different bridge locations, the same four links comprise the top four positions in each case, 37 and 38, and 75 and 76, which correspond to two different directions of the same bi-directional links. These links are shown with red dotted lines on Figure 5. Links 15 and 16, and 23 and 24, which also correspond to two directions of the same bi-directional links, are also important, because they occupy the 5th highest ranked CPIM values. These are shown with purple dashed lines in Figure 5. Based on these results, it is clear that merely retrofitting only one bridge class throughout the network does not have a large benefit for network behavior. Furthermore, the bridges on bi-directional links (corresponding to 37 and 38, 75 and 76, 15 and 16, and 23 and 24) should be retrofitted/inspected in that order.

3.3 MULTI-SCALE SEISMIC RELIABILITY ANALYSIS OF LARGE INFRASTRUCTURE NETWORKS USING HIERARCHICAL STRUCTURES

Using the earlier experimental method, the five AHSIA obtain several bounds results, as shown in Table 5. Two levels of bounds are considered: within 5% and within "closest %". "Closest %" bounds are either within 1% bounds or the closest bounds that can be obtained when all link-sets and cut-sets are accounted for. Note that this table also contains the 5% c.o.v. crude-MCS estimate for comparative accuracy. Of these AHSIA, the RI-AHSIA and the N-AHSIAs perform best in terms of probability; however, RI-AHSIA requires the least amount of time. While the I-AHSIAs return the fastest results, they vastly underestimate the disconnection probability.

Table 5. US Pipeline Network AHSIA Comparison.

Algorithm (MCS: 67.0%, 5% c.o.v.)		Closest % bounds		5% bounds		Time (sec)
		LB	UB	LB	UB	
I-AHSIA	Bi	49.4	50.3	46.5	50.8	2,434
	Min	47.7	48.6	44.5	49.3	2,378
RI-AHSIA		65.7	66.5	63.8	68.5	2,874
N-AHSIA	Stage	72.0	73.0	71.8	74.8	3,817
	Global	71.0	72.0	70.9	73.9	3,628

For an example of the hierarchical structure of this network, see Figure 11, which shows the corresponding RI-AHSIA hierarchical structure. Both of the I-AHSIA create the same dendograms, but require one more level than the RI-AHSIA. This additional depth creates too much error. Therefore, RI-AHSIA presents the best performance for this study. Figure 6 shows the ten components with the highest CPIM values using the RI-AHSIA structure, while Table 6 lists their CPIM values. The top six ranked components are the unidirectional links, i.e., the top 3 bi-directional links, closest to the epicenter. The other four links correspond to the unidirectional links on the other side of the fault. For this example, the top ranked CPIMs are strongly driven by the epicentral distance. It is interesting to note that none of these correspond to the RI-AHSIA intercluster links.

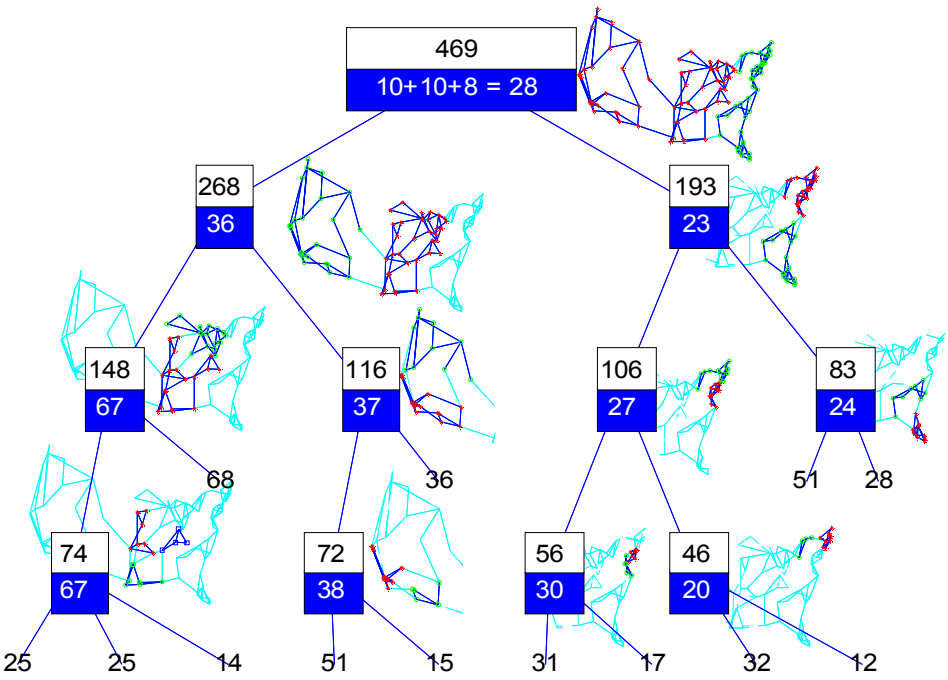

Figure 11. US pipeline network RI-AHSIA hierarchical structure.

Table 6. US Pipeline Network RI-AHSIA Top 10 CPIMs.

Component	Cut-set(%)	Link-set(%)
132	99.4	99.6
133	99.4	99.6
285	99.3	99.4
284	99.3	99.4
131	97.3	97.4
130	97.3	97.4
460	97.0	97.2
461	97.0	97.2
274	90.1	90.3
275	90.1	90.3

3.4 GIMS

The refined bridge network similar to LA is analyzed using the aforementioned experimental approach. The multi-scale S-RDA analysis resulted in disconnection probabilities of 64.4%-65.3% using 1% bounds, and 62.2%-67.1% using 5% bounds, while crude-MCS resulted in a disconnection probability of 66.8% with 5% c.o.v. The bounds definitions have the same

explanation as discussed in Section 3.3, except, since the bounds all converge well, 1% is used instead of closest%. This result is well-contained by the 5% bounds, but not within the 1%, indicating that some error, which was also present in the earlier analysis, is present. See Figure 12 for the RI-AHSIA hierarchical structure. This structure indicates how bridge management might be organized for this network. When organization is not dictated by geo-political boundaries, management bodies may classify bridges into lowest level groupings based on the relevant links. Furthermore, higher level managers may manage those at the leaf nodes based on the locations of junctures in Figure 12. The highest level manager corresponds to the juncture with numerator 311.

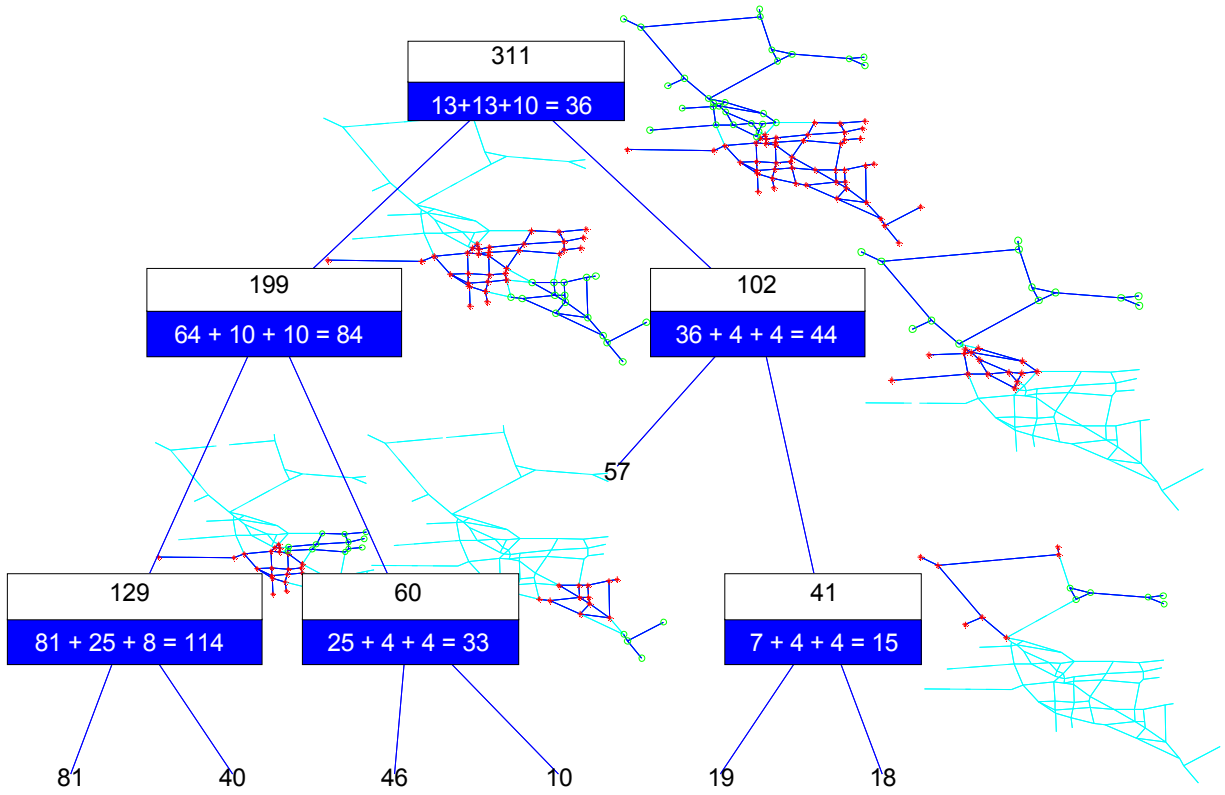

Figure 12. Bridge Network RI-AHSIA Hierarchical Structure.

For a description of component importance, Figure 7 shows the components having the top CPIM values, while Table 7 lists the values. Note that there seems to be coarse convergence throughout the table, but especially for the top ranked components. Only two of these components, 112 and 113, correspond to the directions of one bi-directional intercluster link, which occur on the lowest level of the dendogram. Overall, the highest ranked links are strongly driven by epicentral distance. Furthermore, 88 and 89 correspond to a bi-directional link that is peripheral, which probably has elevated importance due to high correlations with more important components. In addition, the cut-set definition has a higher value than the link-set definition, indicating poor convergence.

Table 7. Bridge Network Top 10 CPIMs.

Component	Cut-set(%)	Link-set(%)
88	79.9	82.3
89	79.8	82.3
112	67.4	69.5
113	67.5	69.5
90	66.7	68.6
91	66.6	68.6
110	59.9	61.8
111	59.9	61.8
246	56.6	58.3
247	56.6	58.3

To overcome such component issues, a regional perspective based on GIM rankings will be investigated. See Figure 13 for a barplot of the GIM values. These clusters are labeled from 0 to the total number of clusters from the first branch on the left to the bottom right leaf cluster in Figure 12, above. Because the bridge component definition uses infinitely reliable nodes, these are omitted from the intersection description. It is clear from the barplot that there is much more variation for the intersection description than for the union description, and that cut-set-based and link-set-based approaches converge well. This indicates that there is more information in the intersection description. In general, the intersection values are smaller.

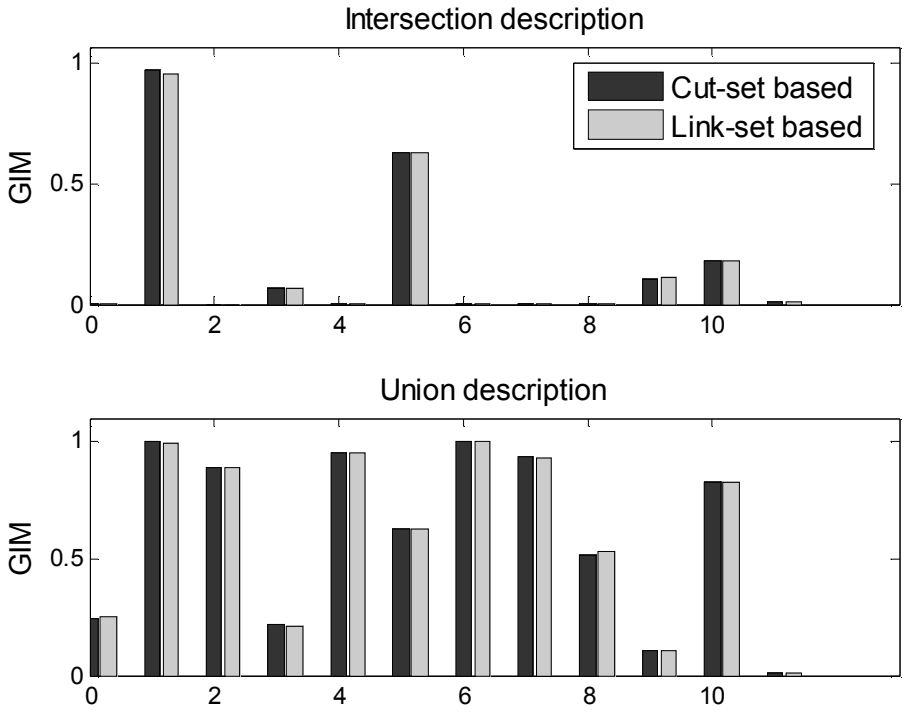

Figure 13. Bridge Network GIM Values.

Because the values are hard to connect to the regions and network resolutions they represent, the link-set based group ordinal rankings in the dendogram structure are shown in Figure 14. The ranks in the tops of the ovals represent the intersection description, while the bottom numbers represent the union description. The link-set definition is used because more link-sets were identified in the multi-scale S-RDA approach. This figure contains a large amount of information; however, the source node's influence is most important. For both descriptions, the branch that ends with the leaf cluster containing the source node, i.e., the cluster ranked 3 and 6 by intersection and union, respectively, with size 19 is highly important. The top of the branch is of size 102, yet still receives the number one ranking from intersection, while the small-sized leaf cluster containing the source node is size 19, while still being ranked highly by union. The regions containing the source node must be given special treatment by management bodies. The leaf cluster containing components 88 and 89, which are the top ranked components in terms of the CPIM, receives an intersection rank of 7, showing that it is not as important. Note that these groupings are, like the components, all close to the epicenter; however, the sub-branch containing the terminal node, which is far from the epicenter, also receives a high intersection ranking of 4, with the cluster containing it being assigned a ranking of 5. The terminal node cluster receives an elevated ranking due to its small size, so it does not need to be given as much consideration by management. These group metrics are less sensitive to the issues that affect the component importance metrics. Both show that several components and groups closest to the epicenter and participating in network function are important.

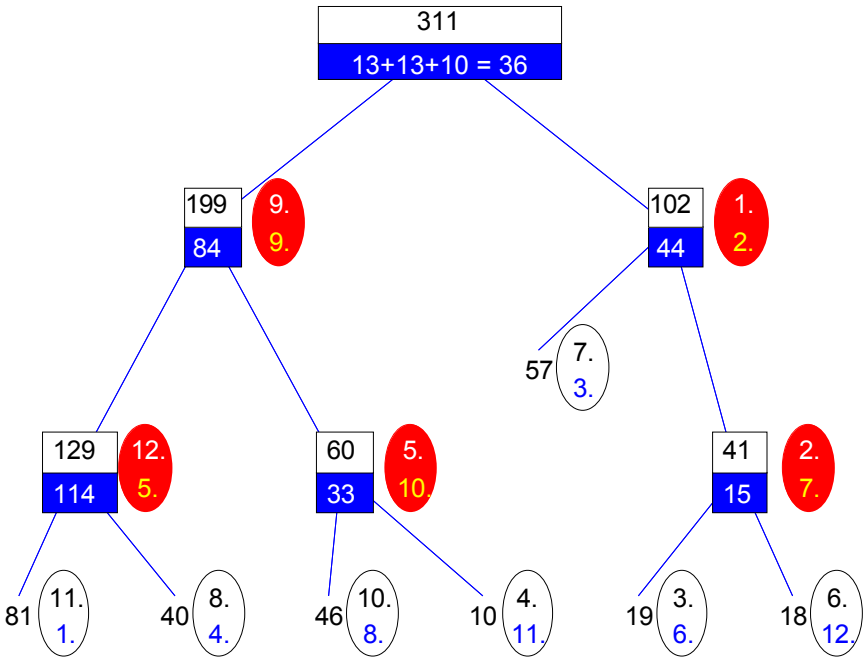

Figure 14. Bridge Network GIM Link-set Rankings.

3.5 DETERIORATING EFFECTS ON PIPELINE NETWORKS

Using the experiment design and parameters discussed earlier, a deteriorating pipeline network will be investigated. First, the difference between the connectivity-based and the reliability-based hierarchical structures will be investigated. Next, the S-RDA bounds for each structure at each time point will be explored. The component CPIMs for each structure will be investigated last.

For a description of the connectivity- and reliability-based hierarchical structures, see Figure 15. While these two hierarchical structures have the same depth, the major difference between them is that the reliability base does not specify a leaf cluster of size 4. This reliability-based structure is also more balanced on the top-level division. Although the reliabilities degrade over time, the reliability-based hierarchical structures remain constant over time. Furthermore, the S-RDA bounds for each time point can be found in Table 8. The bounds descriptions are the same as provided in the Section 3.3. The Time column corresponds to the amount of time (in seconds) required to obtain the closest% bounds. While the connectivity-based S-RDA bounds contain the 5% c.o.v. crude-MCS bounds best, the reliability-based S-RDA bounds exhibit better convergence. This stronger convergence indicates further why the reliability-based S-RDA bounds require more time. Because the connectivity-based S-RDA at 13 years cannot converge within 5%, coarse convergence is seen for the connectivity-based structure. The results for this case are listed as not applicable (NA). Because the reliability-based approach underestimates when compared to crude-MCS, for two out of three cases, the connectivity-based approach contains the crude-MCS estimate best, indicating that it is the most accurate. However, the reliability-based approach is better for the 13-years case, because it converges much better.

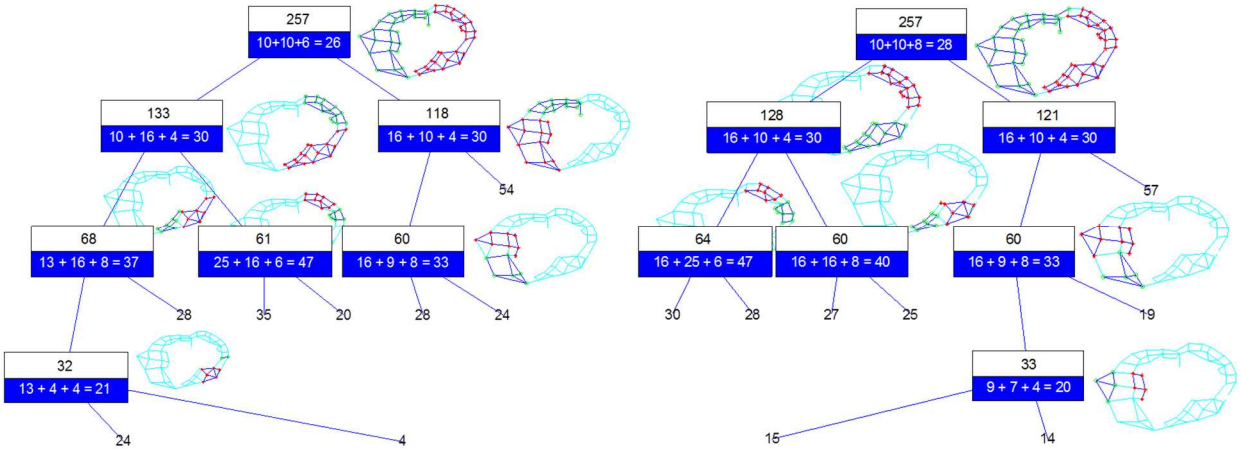

Figure 15. Pipeline Network Hierarchical Structures.

Table 8. S-RDA Bounds.

t (yrs)	Crude-MCS 5% c.o.v.	Connectivity					Reliability				
		5%		Closest%		Time (sec)	5%		Closest%		Time (sec)
		LB	UB	LB	UB		LB	UB	LB	UB	
0	52.3	47.5	52.5	51.5	52.5	436	44.6	49.5	48.4	49.4	538
13	52.3	NA	NA	48.0	54.1	702	47.5	52.4	49.0	52.3	948
25	60.5	55.1	60.0	57.8	60.0	635	54.1	59.0	56.6	58.9	652

For the 0 year case, the highest ranked CPIMs values are shown in Table 9, below, while their locations are listed in Figure 9. Note that two of these components are nodes, with one being the terminal node. Also note that both structures rank the top CPIMs the same, with very similar values.

Table 9. 0 Year Pipeline Top 10 CPIMs.

	Connectivity		Reliability	
Component	Cut-set(%)	Link-set(%)	Cut-set(%)	Link-set(%)
110	39.0	40.7	39.6	40.4
111	39.0	40.7	39.6	40.4
102	36.0	37.6	36.6	37.3
103	36.0	37.6.	36.6	37.3
256	34.6	36.1	35.0	35.7
257	34.6	36.1	35.0	35.7
97	34.2	35.8	34.7	35.4
96	34.2	35.8	34.7	35.4
54	32.3	34.3	34.2	35.2
46	30.8	32.6	33.8	34.7

The highest components for the 13-year and 25-year deterioration cases are shown in Figure 16. For either the connectivity- or reliability-based approach, the highest ranked components, by CPIM values, that occur only for 13 years of deterioration are shown in gold. The highest ranked components that occur only for 25 years of deterioration are shown in black. The top ranked components for both deterioration scenarios are shown in grey. The corresponding top ranked CPIM values for both structures are shown in Table 10 and Table 11, for 13 years and 25 years of deterioration, respectively. Note that the connectivity- and reliability-based top ranked CPIM differed in both the order of rankings and in the components specified for the 13-year deterioration case. Recall the coarseness of the connectivity-based S-RDA bounds for the 13-year deterioration case. This results in more discrepancy between the different structures.

Lastly, the 25-year deterioration case ranks the same top 10 components, but orders them differently. These bounds are less coarse than the 13-year deterioration case. Furthermore, it is evident from the location description in Figure 16 that the top ranked CPIM components eventually all become nodes near the seashore; however, the nodes closest to the seashore are not given the highest rankings, and "38" and "28" do not appear in the top 10 for the 25-year deterioration case. While the rankings are sensitive to the node fragility fitted FIFs, this indicates that network function still strongly influences which components are most important. The deterioration effects seem to stabilize as more deterioration occurs. Furthermore, the deterioration has larger system effects later in the analysis time, while more effects happen at the component level earlier on.

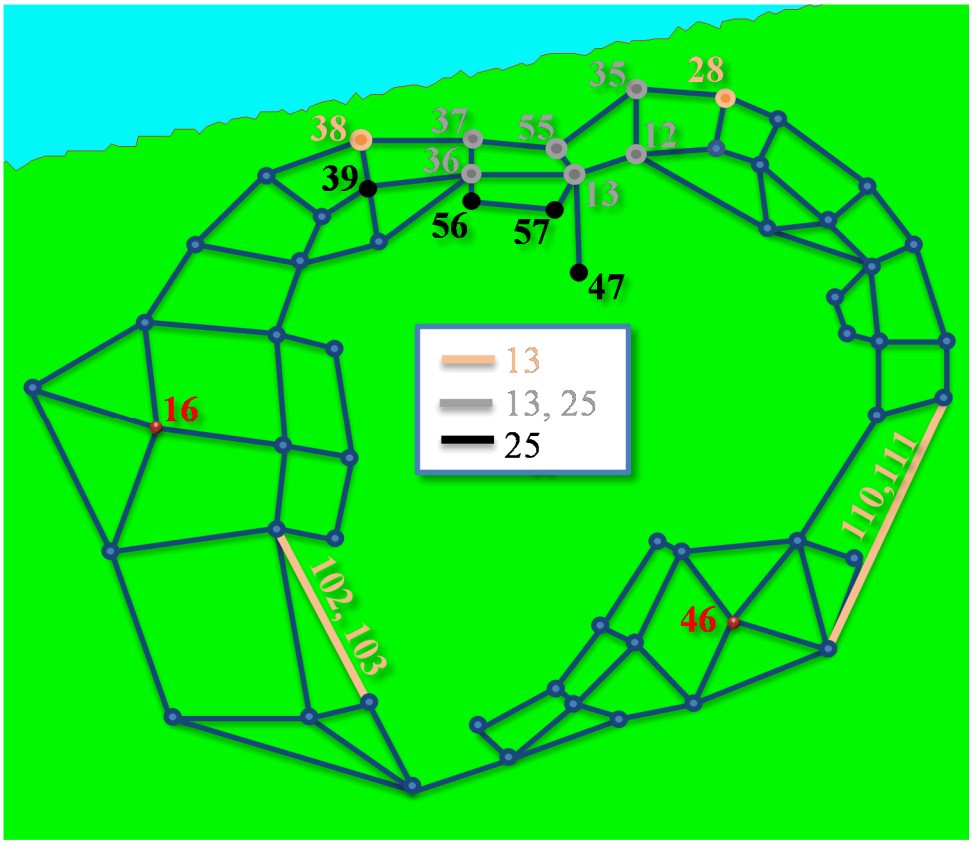

Figure 16. Pipeline network hierarchical structures.

Table 10. 13-Year Pipeline Top 12 CPIMs.

	Connectivity		Reliability	
Component	Cut-set(%)	Link-set(%)	Cut-set(%)	Link-set(%)
35	51.5	45.6	51.6	49.7
55	51.1	45.3	34.8	33.6
38	45.9	33.0	42.3	28.8
37	45.8	23.7	41.5	36.6
13	42.3	40.3	41.7	41.7
12	42.0	39.7	41.0	41.2
36	41.0	33.6	35.6	34.4
110	40.1	40.3	39.9	40.2
111	40.1	40.2	39.9	40.2
102	37.0	37.1	36.7	37.1
103	36.9	37.1	36.7	37.1
28	33.9	29.1	39.5	39.3

Table 11. 25-Year Pipeline Top 10 CPIMs.

	Connectivity		Reliability	
Component	Cut-set(%)	Link-set(%)	Cut-set(%)	Link-set(%)
55	79.6	83.3	69.4	68.0
13	78.6	81.7	80.2	72.7
12	75.1	77.2	76.7	73.6
35	74.1	74.2	67.6	72.3
36	66.8	70.4	70.7	70.1
56	66.5	69.9	70.3	69.5
37	66.4	69.7	70.3	69.3
57	65.1	67.9	69.3	67.5
47	63.7	65.6	67.6	65.3
39	63.2	65.0	67.4	64.6

4. DISCUSSION

4.1 CE-BASED ADAPTIVE IS USING GAUSSIAN MIXTURE

Using the proposed method for adaptive IS to address a series system showed the value of using a multimodal distribution model, and that the near-optimal IS density could be found using fewer iterations of the algorithm. Furthermore, in Kurtz and Song (2013) this approach was found to be 1) efficient and accurate to the level of both probability and limit-state curvatures, and 2) to perform well for a variety of component and system problems, including series and general systems. Additionally, noisy limit states exhibited few problems for this approach. Lastly, it is apparent that, if the distribution of experimental observations and the prior distributions are supplied, a Gaussian mixture could be used to fit the unknown optimal distribution by minimizing the CE distance between it and the overlap.

4.2 SEISMIC RELIABILITY ANALYSIS OF DETERIORATING REPRESENTATIVE US WEST COAST BRIDGE TRANSPORTATION NETWORKS

Using the multi-scale, time-variant, seismic, bridge model discussed earlier, several bridge networks with the same topology, but at different bridge locations, were analyzed at different stages of deterioration using S-RDA bounds and CPIMs. The use of this approach provided a way to apply the component-driven network model in Figure 1 to bridge networks. Using accurate probabilistic capacity and demand models, along with application-specific bridge classes, a transportable methodology that does not over-generalize allows for application-specific results, and presents a significant benefit over existing analysis approaches. The earlier example demonstrated the dependence on the bridge configuration, the amount of deterioration, and location of the epicenter. Fortunately, it also indicated that the most important links needed to improve via bridge retrofit or repair were not strongly deterioration dependent. In Kurtz, et al. (under review), tests addressing the effects of time-variance for the same bridge locations, spatial correlations, and subjunctive representations were investigated. Furthermore, this study found that the approximations used for efficiency purposes did not introduce intolerable error.

4.3 MULTI-SCALE SEISMIC RELIABILITY ANALYSIS OF LARGE INFRASTRUCTURE NETWORKS USING HIERARCHICAL STRUCTURES

Several of the AHSIAs used to obtain hierarchical structures for S-RDA were used for earlier investigations that included a possible realization of a US-wide pipeline network subjected to one of the New Madrid seismic events of the early 1800s. The results of this analysis showed that the components with the highest valued CPIMs were most sensitive to the location of the epicenter, and that RI-AHSIA performed the best. The study in Kurtz (under review) further investigated a regular network, a starlike topology (as found in transmission networks), a large network, and a more refined version of the earlier-used bridge network. These studies further

suggest that the RI-AHSIA is the most robust and the most appropriate to use with S-RDA. Furthermore, the hierarchical structure using RI-AHSIA identifies the most intuitive form.

4.4 GIMS

GIM formulations using union and intersection descriptions for both link-set and cut-set definitions were developed and tested on a more refined version of the bridge network. As was found with the previous bridge network study, the epicentral distance again determined many of the groups found to be most important. Specifically for the GIMs, the location of the source node strongly affected which groups were most important. In the hierarchical structures, those containing highly ranked clusters were also found to be highly ranked themselves. These structures also determine how management should be organized at several levels of resolution. In Kurtz (under review), a pipeline network with a clear hierarchical structure was also analyzed. This network seemed to indicate that the most important groups of components contained neither the source nodes nor the terminal nodes, but did contain the groups that were most important to traversing between the two. Finally, from these analyses it was found that, even when the intersection description has more information content due to the larger variability in values, both descriptions are necessary for determining which groups are most important. Furthermore, the larger groups with high intersection description ranking, as well as the smaller groups with high union description ranking, must be given special attention by management organizations.

4.5 DETERIORATING EFFECTS ON PIPELINE NETWORKS

A deteriorating pipeline network was investigated to further determine the effects of deterioration on several network rubrics. The component level was most sensitive to the deterioration, while the system level showed smaller changes. Particularly for the connectivity-based analysis, the network analysis convergence was also strongly sensitive to the level of deterioration. In Kurtz (under review), this same network was re-evaluated using only link-based deterioration. This approach resulted in changes to the reliability-based hierarchical structure with time. For these pipeline networks, the connectivity-based approach showed less sensitivity, but more accuracy, with deterioration. A bridge network with non-uniform deterioration was also investigated. While the bridge model exhibited increases in values for all network rubrics, the ordinal rankings exhibited very little change. The bridge network example showed that the reliability-based approach was most accurate. This indicates that pipeline networks were far more sensitive to deterioration than the bridge network.

5. ANTICIPATED IMPACT

This report has focused on several ways to perform adaptive sampling and analysis of deteriorating infrastructure networks with a focus on group metrics and hierarchical structures. As is the case for most research, there are always improvements to be considered. The adaptive IS approach 1) is able to solve all structural reliability problems for both component and system examples, 2) is not sensitive to the level of probability and limit state curvatures, 3) can handle multiple design point problems, and 4) can handle limit states with numerical noise. This analysis also allows the user to identify the most important regions of the failure domain.

A simple advancement for this method would be to investigate a Bayesian parameter estimation example directly. While the adaptive IS approach presents a significant benefit over general IS, the computational complexity can be further decreased by making the number of densities during the search adaptive, based on the densities' relative importance, and re-using the pre-samples in the final IS. Lastly, this approach can be further applied to even higher complexity applications, e.g., deterioration parameter maps and bridge parameters. Such network applications also show rather interesting uses of existing data. For example, using strategic field samples of surface chloride concentrations, deterioration maps may be updated at uninspected areas, presenting even more future benefit.

These deterioration maps are of specific interest to the bridge component-driven network model. This component model enables the analysis of the general transportation networks that are predominantly governed by bridge failures in seismic zones of large scale networks, e.g., several thousand bridges; however, this network can be further expanded to failures of other structures, e.g., fault surface ruptures of pavement, tunnel failures, culvert failures, landslides of embankments, etc. Furthermore, this model can be applied to an exact topology of an existing network with accurate bridge locations to suggest optimal management strategies to an existing agency. Because this approach was developed for major state highways, the resolution of this network must be given special consideration. The node models may also be advanced to use the behavior of representative highway bridge interchanges, which use bridge structures that are outside the parameter range of current experimental data. These suggestions are not overly specific to bridge networks, but also have broader impacts on general infrastructure networks.

Indeed, general infrastructure network analysis can be enhanced using hierarchical structures to approach large-scale complex networks. The automated multi-scale network analysis approach presents significant computational efficiency improvements for the use of S-RDA for networks exhibiting a multi-scale cluster structure; however, this approach ought not to be used for regular or scale-free networks. It should be noted that such a hierarchical approach is completely novel to civil infrastructure network analysis, and presents significant benefit. However, as a network disconnection analysis, this approach uses specific source and terminal node pairings, and should

not be used for subjunctive representations, as these inherently change the cluster structure. The source and terminal nodes correspond to those evacuation/service corridors that are most critical. To generate a hierarchical structure for management, these should be placed so that they affect the hierarchical structure the least. Furthermore, the super-link representation can be made more efficient by placing a "super-node" inside the cluster. This provides the added benefit of using a linear rule to determine the number of super-links, as opposed to using a quadratic rule. Lastly, this approach ought to be tested on multimodal networks, e.g., water pipeline and electric transmission networks with connections at pump stations, using varying degrees of intermodal connectivity. The majority of network analyses neglect such interdependency, and, depending on the level of mode interaction, may be largely inaccurate. Further study of such multimodal networks is absolutely necessary.

Because many of the network outputs are sensitive to the hazard model, the most general model ought to be used. The hazard model currently relies on attenuation rules from the Chi-Chi earthquake records in Taiwan. These can be further improved for specific applications by the use of local data. Furthermore, the seismicity of an exact earthquake can be generalized by using hazard maps. Fortunately, the current methodology is fully capable of using this information. Both component-level importance and, to a lesser extent, the group-level importance are affected by an exact earthquake. Hazard maps will give more general information about network management strategies. As is necessary for all proposed numerical frameworks, a trial of an exact network application, possibly in coordination with a utility company or a state Department of Transportation, is required to suggest specific management solutions.

While the hazard model is a readily available input with typically low computational expense, the compiling of the component model is very expensive, computationally, for large infrastructure networks. This is especially true for the correlation terms, as they vary quadratically with the number of components. While parallelization may solve this issue, powerful computational resources may not always be available. With this end in mind, a statistical learning method may be able to estimate the majority of correlation terms without the need for additional expensive numerical integrations using a subset of correlation terms. The exact estimation procedure will probably be network specific, making such models necessary for each application. As a tentative approach, a regression, based on component inputs, that uses a stepwise elimination to create a parsimonious model, may be used. Data transformations to observe homoscedasticity and a division of data for fitting may also be required. Note that the proposed component model is not computationally expensive for small networks.

Lastly, the deterioration model, which is easily handled by the current methodology, must be updated in several ways. An analysis of deteriorating component sections should be attempted specifically for pipeline models. The current approach is purely theoretical. Furthermore, the deterioration maps should be based on field data. A study focusing on how de-icing salts are

distributed is also necessary. More general coastlines must be investigated, i.e., bays, islands, peninsulas, isthmuses, etc. While the current methodology allows for a high degree of flexibility, the current list of potential benefits, border impacts, and improvements, as well as the need for further study is by no means exhausted.

6. CONCLUSION

This study has sought to address many issues relevant to infrastructure network analysis: (1) adaptive use of new information, (2) extending analytical, component-driven approaches based on structural reliability fundamentals to deteriorating bridge networks, (3) AHSIA with S-RDA, (4) the use of hierarchical structures for GIMs, and (5) deterioration effects. The adaptive IS approach has been found to be useful for a wide range of structural reliability problems. The deteriorating bridge network analysis was able to determine that, for the given network, the deterioration did not heavily affect which links ought to be serviced first. The AHSIA, particularly RI-AHSIA, along with S-RDA, presented significant reductions in computational complexity for certain network types, especially those which exhibited clusters. The hierarchical structures and GIMs were also able to suggest regional organization and management strategies. Lastly, the deterioration study described a way in which seaside deterioration may affect a city's pipeline network.

Throughout this study, various network metrics were developed to create a risk-based framework that could be used to recommend optimal management strategies to decision makers. These suggestions extend to several resolutions: the component level, the group level, the hierarchical organization, and the likelihood of network disconnection. Depending on what decision makers principally desire, these suggestions can become more specific. Without such suggestions, the nature of the group management approach becomes somewhat unbounded.

While each of these points present advances in several types of knowledge, they can be further improved, especially for the more tentative studies. The approaches are applicable to existing, complex networks, and must be demonstrated on exact networks in conjunction with the participation of an appropriate management body. Such an application will require several uses of specific network data. Furthermore, this framework is general enough to analyze many kinds of civil infrastructure networks that are subject to seismicity.

REFERENCES

1. ASCE, *2013 Report Card for America's Infrastructure*, The American Society of Civil Engineers, 2013.
2. Ang, G.L., Ang, A.H.S., and Tang, W.H. *Multidimensional kernel method in importance sampling*, Proceedings of the 6th International Conference of Applications of Statistics and Probability in Civil Engineering, Eds. Esteva, L., and Ruiz, S.E. 1: 289-296, 1991.
3. Box, G.E.P., and Tiao, G.C. *Bayesian inference in statistical analysis*, John Wiley and Sons, Inc., Hoboken, NJ. 1992.
4. Bucher, C.G. *Adaptive sampling – an iterative fast Monte Carlo procedure*, Structural Safety. 5: 119-126, 1988.
5. CalTrans. *California Log of Bridges on State Highways*. CalTrans, Structure Maintenance and Investigations. 2013.
6. Ching, J., and Hsieh, Y. *Local estimation of failure probability function and its confidence interval with maximum entropy principle*, Probabilistic Engineering Mechanics. 22(1): 39-49, 2006.
7. Choe, D., Gardoni, P., and Rosowsky, D. *Closed-form fragility estimates, parameter sensitivity and Bayesian updating for RC columns*, Journal of Engineering Mechanics. 133(7): 833-843, 2007.
8. Der Kiureghian, A. *First- and Second-order Reliability Methods*, Engineering Design Reliability Handbook, Eds. Nikolaidis, E., Ghiocel, D.M., and Singhal, S. CRC Press, Boca Raton, FL., Chap. 14, 2005.
9. Der Kureghian, A., and Dakessian, T. *Multiple design points in first and second-order reliability*, Structural Safety. 20(1): 37-49, 1998.
10. Der Kiureghian, A., and Song, J. *Multi-scale reliability analysis and updating of complex system by use of linear programming*, Reliability Engineering and System Safety. 93(2): 288-297, 2008.
11. Dubourg, V., Deheeger, F., and Sudret, B. *Metamodel-based importance sampling for simulation of rare events*, in Applications of statistics and probability in civil engineering, Proceedings of the 11th International Conference on Applications of Statistics and Probability in Civil Engineering, Eds. Faber, H., Kohler, J., and Nishima, K. 661-668, 2011.
12. Elnashai, A.S., and Di Sarno, L. *Fundamentals of Earthquake Engineering*, Wiley, Chichester, U.K. 2008.
13. Engelund, S., and Rackwitz, R. *A benchmark study on importance sampling techniques in structural reliability*, Structural Safety. 12: 255-276, 1993.
14. Fu, G., and Moses, F. *Importance sampling in structural system reliability analysis*, in probabilistic Methods in Civil Engineering, Proccedings of the 5th ASCE Specialty Conference, Ed. Spanos, P.D. 340-343.
15. Fujita, M., and Rackwitz, R. *First- and second-order reliability estimates by importance sampling*, Structural Engineering/Earthquake Engineering. 5: 53-59, 1988.
16. Gardoni, P., Der Kiureghian, A., and Mosalam, K. *Probabilistic Capacity Model and Fragility Estimates for Reinforced Concrete Columns Based on Experimental Observations*, Journal of Engineering Mechanics. 128(10): 1024-1038, 2002.
17. Gardoni, P., Mosalam, K., and Der Kiureghian, A. *Probabilistic Seismic Demand Models and Fragility Estimates for RC Bridges*, Journal of Earthquake Engineering. 7(1): 79-106, 2003.

18. Gardoni, P., and Rosowsky, D. *Seismic Fragility Increment Functions for Deteriorating Reinforced Concrete Bridges*, Structure and Infrastructure Engineering. 9(1): 969-978, 2009.
19. Genz, A. *Numerical computation of multivariate normal probabilities*, Journal of Computational and Graphical Statistics. 1(2): 141-149, 1992.
20. Gomez, C., Sanchez-Silva, M., Dueñas-Osorio, L., and Rosowsky, D. *Heirarchical infrastructure network representation methods for risk-based decision-making*, Structure and Infrastructure Engineering: Maintenance, Management, Life-cycle Design and Performance. 9(3): 260-274, 2013.
21. Grooteman, F. *An adaptive importance sampling method for structural reliability*, Probabilistic Engineering Mechanics. 26(2): 134-141, 2011.
22. Guikema, S., and Gardoni, P. *Reliabiity Estimation for Networks of Reinforced Concrete Bridges*, Journal of Infrastructure Systems. 15(2): 61-69, 2009.
23. Hastie, T., Tibshirani, R., and Friedman, J. *Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, Inc. New York, New York. 2009.
24. Holt, J., and Chung, S. *Efficient Mining of Association Rules in Text Databases*, Proceedings on 8th Conference on Information and Knowledge Management. 234-242, 1999.
25. Huang, Q., Gardoni, P., and Hurlebaus, S. *Probabilistic Seismic Demand Models and Fragility Estimates for Reinforced Concrete Highway Bridges with one Single-Column Bent*, Journal of Engineering Mechanics. 136(11): 1340-1353, 2010.
26. Huang, Z. *Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values*, Data Mining and Knowledge Discovery, Kluwer Academic Publishers, Netherlands, 2: 283-304.
27. Kaufman, L., and Rousseeuw, P.J. *Finding Groups in Data: an Introduction to Cluster Analysis*, Structural Safety. 32(1): 35-41, 2010.
28. Kurtz, N. *New methodologies for multi-scale time-variant reliability analysis of complex lifeline network*, Ph.D. thesis, The University of Illinois at Urbana-Champaign, Urbana, IL, USA. Under review.
29. Kurtz, N., and Song, J. *Cross-Entropy-Based Adaptive Importance Sampling Using Gaussian Mixture*, Structural Safety. 42: 35-44, 2013.
30. Kurtz, N., Song. J., and Gardoni, P. *Seismic Reliability Analysis of Deteriorating Representative US West Coast Bridge Transportation Networks*, Special Issue on *Resilienc-based design of structure and infrastructures*, Journal of Structural Engineering. Under review.
31. Laarhoven, T., and Marchiori, E. *Graph clustering with local search optimization: The resolution bias of the objective function matters most*, Physical Review E. 87, 012812. 2013.
32. Lee, Y.J., Song, J., Gardoni, P., and Lim, H.W. *Post-hazard flow capacity of bridge transportation network considering structural deterioration of bridges*, Structural and Infrastructure Engineering: Maintenance, Management, Life-cycle Design and Performance. 7(7): 509-521, 2010.
33. Lim, H.W., and Song, J. *Efficient Risk Assessment of Lifeline Networks Under Spatially Correlation Ground Motions Using Selective Recursive Decomposition Algorithm*, Earthquake Engineering and Structural Dynamics. 41(13): 1861-1882, 2012.

34. Lim, H.W., Song, J., and Kurtz, N. *Seismic Risk Assessment of Lifeline Network Using Clustering-based Multi-scale Approach*, Earthquake Engineering and Structural Dynamics. In print (DOI: 10.1002/eqe.2472).
35. Liu, M., and Frangopol, D. *Time dependent bridge network reliability: novel approach*, Journal of Structural Engineering. 131(2): 329-337, 2005.
36. Liu, P.L., and Der Kiureghian, A. *Optimization algorithms for structural reliability*, Structural Safety. 9(3): 161-177, 1991.
37. Mahoney, J. *Correlation Products and Risk Management Issues*, FRBNY Economic Policy Review. New York, NY. 7-20, 1995.
38. Melchers, R.E. *Importance sampling in structural systems*, Structural Safety. 6: 3-10, 1989.
39. Melchers, R.E. *Structural reliability analysis and prediction*, John Wiley and Sons, Inc., Hoboken, NJ. 1999.
40. Melchers, R.E., and Ahammed, M. *Estimation of failure probabilities for intersections of non-linear limit states*, Structural Safety. 23(2): 123-135, 2001.
41. Opsahl, Tore. *Why Anchorage is not (that) important: Binary ties and Sample selection*. http://toreopsahl.com/2011/08/12/why-anchorage-is-not-that-important-binary-ties-and-sample-selection/. 2011.
42. Rokneddin, K., Ghosh, J., Dueñas-Osorio, L., and Padgett, J. *Bridge retrofit prioritisation for ageing transportation networks subject to seismic hazards*, Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance. 9(10): 1-17, 2013.
43. Rubinstein, R. *Simulation and the Monte Carlo methods*, John Wiley and Sons, Inc., Hoboken, NJ. 1981.
44. Rubinstein, R., and Kroese, D. *The cross entropy method: a unified approach to combinatorial optimization, Monte-Carlo Simulation, and machine learning*, Springer, Inc., New York, NY. 2004.
45. Shinozuka, M. *Basic analysis of structural safety*, Journal of Structural Engineering. 109: 721-740, 1983.
46. Song, J., Kang, W.H., Kim, K.S., and Jung, S. *Probabilistic shear strength models for reinforced concrete beams by Bayesian updating based on experimental observations*, The 5th Computational Stochastic Mechanics Conference, Rodos, Greece. 2006.
47. Tibshirani, R., Walther, G., and Hastie, T. *Estimating the number of clusters in a dataset vie the gap statistic*, Technical Report 108, Department of Statistics, Stanford University.
48. Von Luxburg, U. *A tutorial on Spectral Clustering*, Statistics and Computing. 17(4), 2007.
49. Waarts, P.H. *Structural reliability using finite element an appraisal of DARS: Directional Adaptive Response Surface Sampling*, Ph.D. thesis, Delft University of Technology, Delft, Delft, the Netherlands. 2000.
50. Zhang, Y., and Der Kiureghian, A. *Two improved algorithms for reliability analysis*. In Reliability and Optimization of Structural Systems. Proceedings of the 6th IFIP WG 7.5 Woring Conference on Reliability of Structural Systems, Eds. Rackwitz, R., Augusti, G., and Borri, A. 297-304, 1995.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.